



# Chapter 1

## Introduction and Motivation

Particle physics is the subject to study the fundamental structure of the universe. It is now based on the theory called the "Standard Model" (SM). It interprets the universe as the composition of tiny particles interacting with each other by the exchange of force carriers (another type of particle). In July 2012, the discovery of Higgs boson made by the ATLAS and CMS collaboration completed SM 50 years after being predicted. By now, it has been deemed as one of the most successful theories in modern physics.

However, there are still some conflicts between the SM and factual results. For example, in the SM, neutrinos are supposed to be massless, but the discovery of neutrino oscillation support the fact that neutrinos are massive, and the SM cannot explain it. New theories are proposed in order to resolve those conflicts, and they indicate the existence of some new particles or the deviation from SM predictions. This thesis is dedicated to the work in search for this kind of new physics.

### 1.1 Standard Model

The SM is a quantum field theory(QFT). In a QFT the universe is filled with different fields, and all fundamental particles (particles without further substructure) are the forms of quantized fields. They make up the matters and also mediate interactions between them, which is the foundation how this universe operates. Those fundamental

particles could be classified into two types: fermions and bosons. Fermions are the matter builders, while bosons are the force carriers exchanged between particles (for both fermions and bosons).

## Fermions

Fermions are quantized from fermionic fields following Dirac-Fermi statistics with half integer spin number,  $\pm \frac{1}{2}$ . Under the statistic characteristic, fermions exclude each other with the same quantum status, which makes them different from bosons.

All fermions have their counter antiparticles which in the SM have opposite charge and chirality. Those fermions are called "Dirac Fermions". They can be presented as Weyl spinors of four components composed of one left-handed spinor and one right handed spinor following the Dirac equation. However, neutrinos, a sub-specie of fermions, have no counter-partner with opposite chirality found<sup>1</sup>, so they are assumed to be "Majorana Fermions": they are their own antiparticle. They could be instead presented as Majorana spinors in Majorana equation.

**Dirac Equation:**  $i\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0$

**Majorana Equation:**  $i\hbar\gamma^\mu\partial_\mu\psi - mc\psi_c = 0$

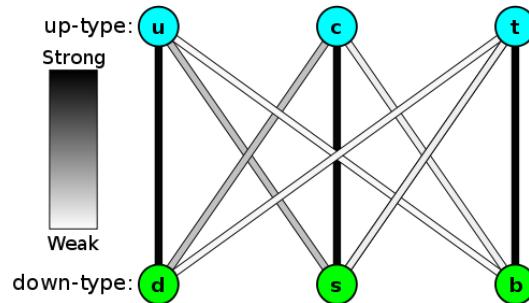
$\psi$  is the fermion field with charge conjugate  $\psi_c$ , and  $\gamma^\mu$  is the gamma matrix and m is the particle mass.

Fermions can then be further categorized into two types, quarks and leptons, by the interactions they participate in. Quarks are the only particles involved in the strong interaction, so they cannot exist alone, and, instead, they are always in bound state as mesons of two quarks or baryons of three quarks.

---

<sup>1</sup>Due to being neutral, although neutrinos and anti-neutrinos were discovered, but neutrinos (anti-neutrinos) only have the left-handed (right-handed) chirality.

Quarks have three generations (flavours) and six flavours. In each generation are quarks with different charges:  $-\frac{1}{3}$  and  $\frac{2}{3}$ . The first generation are the lightest: up and down. Strange and charm are in the second generation. The third generation has bottom and top with highest mass. Quarks can change their flavour through CKM matrix relating to the weak interaction. The decay relation between quarks is shown in Fig.1.1



**Figure 1.1:** Relation between quarks are determined by CKM matrix.

Similar to quarks, leptons also have 3 generations and 6 flavours. In each generation, there is one neutral neutrino and corresponding charged particle with charge -1. The three generations are electrons, muons and taus with their neutrino partners (among them, electron neutrino is assumed to be the lightest fundamental particle under normal hierarchy). Leptons participate in weak interaction, quantum electrodynamics(except for neutral neutrinos) and gravity. They can change flavours through PMNS matrix relating to the weak interaction.

## Interaction and Bosons

Under the SM, the interactions between particles are induced by gauge fields which could be quantized into gauge bosons. Different from fermions, those bosons follow Bose-Einstein statistics with integer spin number, which means more than one is allowed to occupy a single quantum state. They mediate interactions between particles including themselves.

Although there are four fundamental forces in the universe, only three of them are in SM, because they are quantizable: electromagnetic, weak and strong interactions. The challenge of quantizing gravity is still not achieved in modern physics. Each interaction has a part in the SM Lagrange formulism.

The electromagnetic interaction is the best known among the four interactions. It is explained by quantum electrodynamics in the SM. The interaction is induced by electromagnetic field between two charged particles with charge as the invariance under  $U(1)$  symmetry, which could be seen as they interchange photons. Because the electromagnetic interaction only occurs between charged particles, photon doesn't interact with neutral particles at the leading order<sup>2</sup>. The coupling constant (interaction strength which determines the possibility of a process occurs) in the interaction is:

$$\alpha_{EM} = \frac{e^2}{4\pi\hbar c} = \frac{1}{137.036\dots} \quad (1.1)$$

with  $e$  as electric charge of electron,  $\hbar$  as reduced Plank constant and  $c$ , the speed of light. Its part in the SM Lagrange could be written as:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu)\psi - F^{\mu\nu}F_{\mu\nu} \quad (1.2)$$

In this equation,  $\gamma^\mu$  is the Gamma matrices,  $\psi$  is the Weyl spinor of spin  $-\frac{1}{2}$  particles and  $\bar{\psi}$  is the Dirac adjoint of  $\psi$ .  $D_\mu = \partial_\mu + ieA_\mu + ieB_\mu$  represents the gauge covariant derivative with  $e$  as the electric charge,  $A_\mu$  as the filed induced by the particle itself and  $B_\mu$  as the field from external source. In the equation,  $F_{\mu\nu}$  is the electromagnetic field tensor.

All the left-handed particles participate in the weak interaction. It is mediated by three different bosons: the  $W^+$ ,  $W^-$  and  $Z^0$  bosons. They are massive gauge bosons which obtain their mass via the electroweak symmetry breaking. The  $W$  boson is the

---

<sup>2</sup>With a loop diagram, it can still be achieved by exchanging charged fermions between photons.

mediator when a particle changes its flavour along with its charge, while Z boson is involved in the neutral current interactions which leave the particles unchanged with only kinematic momentum transfer. Within this process, a quantity, isospin, is conserved under  $SU_L(2)$  symmetry. Its definition is similar to the spin numbers of a pair of electrons in the same orbital. For two fermions in the same generations, their quantum states are identical except for isospin which is opposite of them to each other. As right-handed fermions don't participate in weak interaction, their isospin is 0. The isospins of fermions are showed in Table 1.1.

**Table 1.1:** Isospin of Elementary fermions

1st Generation	Isospin	2nd Generation	Isospin	3rd Generation	Isospin
$e^-$	$-\frac{1}{2}$	$\mu^-$	$-\frac{1}{2}$	$\tau$	$-\frac{1}{2}$
$\nu_e$	$\frac{1}{2}$	$\nu_\mu$	$\frac{1}{2}$	$\nu_\tau$	$\frac{1}{2}$
$u$	$\frac{1}{2}$	$c$	$\frac{1}{2}$	$t$	$\frac{1}{2}$
$d$	$-\frac{1}{2}$	$s$	$-\frac{1}{2}$	$b$	$-\frac{1}{2}$

The coupling constant for weak interaction is defined as:

$$\alpha_W = \frac{g_W}{4\pi\hbar c} \approx \frac{1}{29} \quad (1.3)$$

with  $g_W$  as the W weak charge strength. In terms of the interactions via Z boson, it is substituted by Z weak charge,  $g_Z$ . A unification between the weak and electromagnetic interactions could be achieved with another new parameter called electroweak hypercharge defined as  $Y_w = 2(Q - I_3)$  where  $I_3$  is isospin and  $Q$  is the electric charge under  $SU_L(2) \times U(1)$  symmetry in the scale of high energy. In SM, the symmetry would be spontaneously broken by the Higgs field to give particles mass. It will be discussed in the next section.

Only quarks are involved in the strong interaction which is described by quantum chromodynamics(QCD). The conserved quantity in the interaction is also imaginary, colour, with gluons as the force carrier boson under  $SU(3)$  symmetry. There are three different colours: red, blue and green along with their anti-colour partners. Similar to

the principle of light, the colour would be absent when the three colours are mixed together or with their anti-colour, and it is the condition for a stable state in QCD. Each quark is only allowed to carry one colour, but this state is unstable. It needs to be bound with another quarks to stabilize the system, and they exchange gluons to form the bounding force. In QCD, gluons have 8 types with different colour combinations:

$$(r\bar{b} + b\bar{r})/\sqrt{2}, \quad -i(r\bar{b} - b\bar{r})/\sqrt{2} \quad (1.4)$$

$$(r\bar{g} + g\bar{r})/\sqrt{2}, \quad -i(r\bar{g} - g\bar{r})/\sqrt{2} \quad (1.5)$$

$$(b\bar{g} + g\bar{b})/\sqrt{2}, \quad -i(b\bar{g} - g\bar{b})/\sqrt{2} \quad (1.6)$$

$$(r\bar{r} - b\bar{b})/\sqrt{2}, \quad (r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6} \quad (1.7)$$

with  $r$ , red charge,  $b$ , blue charge, and  $g$ , green charge.

Different from the other interactions, the colour confinement of gluon self-interaction makes the effective potential increase linearly with the distance between two colour-charged particles. Under this process, the potential energy decays into a quark-antiquark pair, and it is repeated also within the newly produced pair. This leads to the divergence of with the perturbative strong coupling constant, and the mathematical technique, “renormalization”, is introduced to solve the problem.

Its part of the SM Lagrange could be shown as:

$$\mathcal{L}_{QCD} = \bar{\psi}(i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} \quad (1.8)$$

with

$$G_{\mu\nu}^a = \partial_\mu A_\mu^a - \partial_\nu A_\nu^a + g f^{abc} A_\mu^b A_\nu^c \quad (1.9)$$

with  $\psi_i$ , the quark field in  $SU(3)$  representation indexed of i,j, ...,  $G_{\mu\nu}^a$ , the gluon field also in  $SU(3)$  representation indexed of a, b... from 1 to 8.  $f^{abc}$  is the structure constant,  $A_\mu$  is the spin 1 gluon filed and  $g = \sqrt{4\pi\alpha_s}$  is the QCD coupling constant.

All the elementary particles with their basic properties are shown in Fig. 1.2. The 3

interactions with their conserved quantities makes SM under the gauge theory with the gauge group with  $U(1) \times SU(2)_L \times SU(3)$  gauge group.

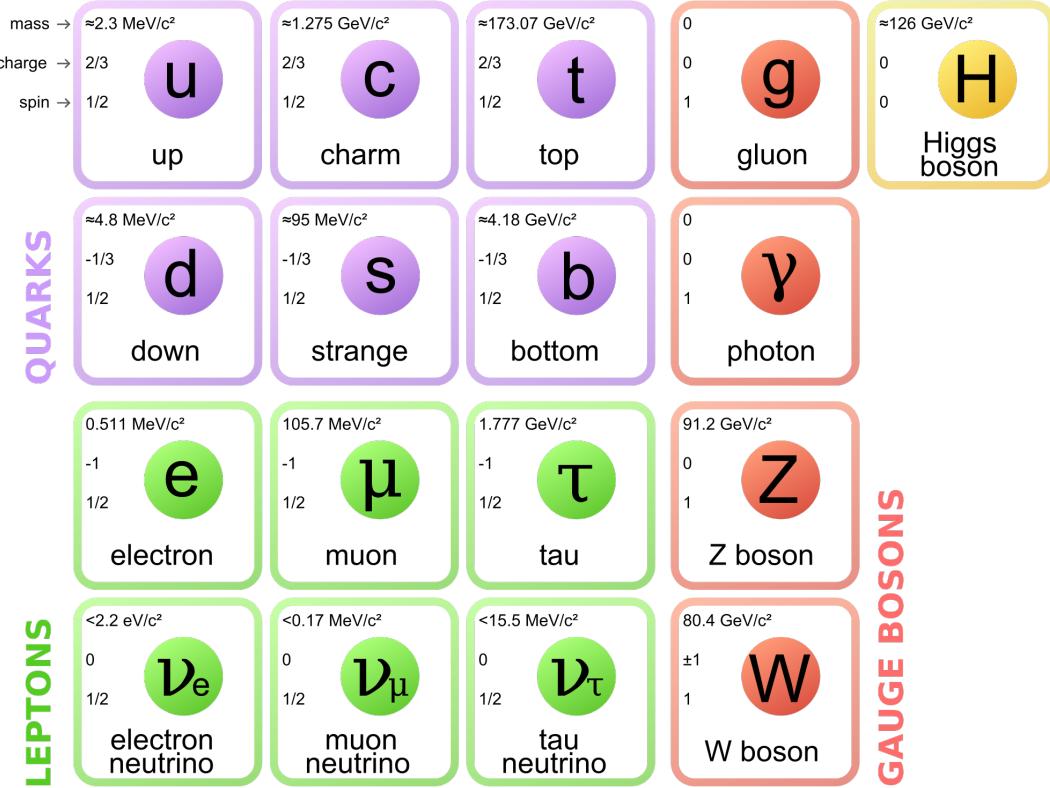


Figure 1.2: Elementary particles and properties

## 1.2 Electroweak Symmetry Breaking

One particle in Fig. 1.2 is not mentioned yet: Higgs boson, the last discovered fundamental particle in the SM. It arises from quantized Higgs field which was proposed by three groups in early 1960s: Robert Brout and Francois Englert, Peter Higgs as well as Gerald Guralnik, C. R. Hagen, and Tom Kibble. It induces spontaneous electroweak symmetry breaking via the "Brout-Englert-Higgs mechanism". The Higgs boson discovery was announced on 4th July 2012 and confirmed on 14 March 2013 with spin 0 and + parity by the ATLAS and CMS collaboration.

The Higgs field is defined as the scalar gauge field in a complex scalar  $SU(2)_L$  doublet

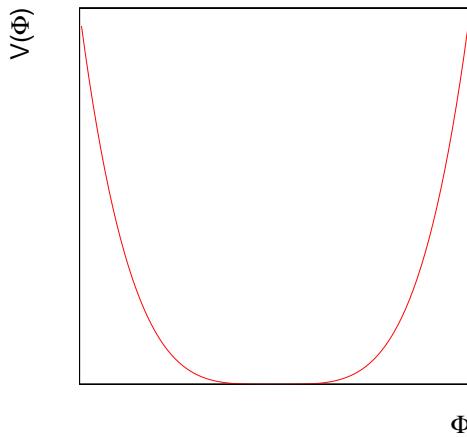
:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.10)$$

The potential for field is:

$$V(\Phi) = \mu^2 |\Phi^\dagger \Phi| + \lambda (|\Phi^\dagger \Phi|)^2 \quad (1.11)$$

For some value of  $\mu$  and  $\lambda$ , the minimal potential can be at  $\Phi = 0$ , and the shape of the potential would be as Fig. 1.3 (this is a simplified plot, and the real one should be in 4 dimensions). In this potential, the symmetry is not broken with the minimal value at  $\Phi = 0$ .

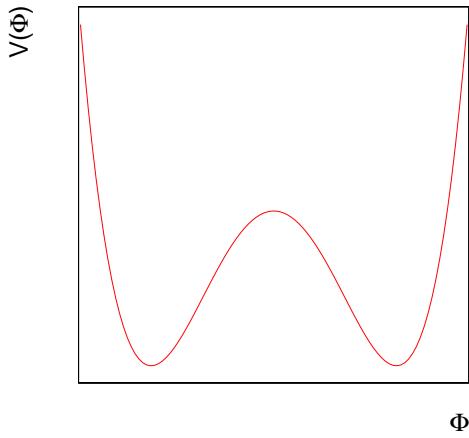


**Figure 1.3:** Scalar potential with  $\mu^2 > 0$

In an alternative scenario for  $\mu^2 < 0$ , the potential shape becomes Fig. 1.4. The minimal expected value of the potential is not at 0 but at:

$$\langle \Phi \rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \frac{\nu}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.12)$$

The potential is only affected by  $|\Phi^* \Phi|$ , so the shape of the potential is determined by the real term (imaginary terms are cancelled out). The value in Eq. 1.12 is called the "vacuum expectation value"(VEV). To maintain a stable state, particles are only allowed to stay in the lowest potential, the valley part. This makes the degree of freedom of the particles decrease from four to one and breaks the  $SU_L(2) \times U(1)$  symmetry with isospin and hypercharge to  $U(1)$  symmetry with electric charge. In high energy



**Figure 1.4:** Scalar potential with  $\mu^2 < 0$

regime above the valley (excited state), electromagnetic and weak interaction are mixed together to form three  $SU_L(2)$  gauge bosons,  $W_\mu^i$  with  $\mu = 1, 2, 3$  and one  $U(1)$  gauge boson,  $B_\mu$ . They are not SM particles, but they could be taken as the excited form of SM gauge bosons. The Lagrangian for the interaction between them and Higgs field is:

$$\mathcal{L} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi) \quad (1.13)$$

with

$$D_\mu = \partial_\mu + i\frac{g}{2}\tau \cdot W_\mu + i\frac{g'}{2}B_\mu Y \quad (1.14)$$

$g$  and  $g'$  are the coupling constants between the fields,  $\tau$  is the Pauli matrix and  $Y$  is the hypercharge.

A unitary gauge transformation on the Higgs field can remove Goldstone bosons<sup>3</sup> after the symmetry breaking. The Higgs field is thus shifted with the new gauge as:

$$\Phi = \frac{\nu + h}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.15)$$

with  $h$ , the physical Higgs sector, as a complex number.

After inserting the new Higgs field into and rearranging SM Lagrange, the SM gauge bosons could be shown as:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (1.16)$$

$$Z^\mu = \frac{-g'B_\mu + gW_\mu^3}{\sqrt{g^2 + g'^2}} \quad (1.17)$$

$$A^\mu = \frac{gB_\mu + g'W_\mu^3}{\sqrt{g'^2 + g^2}} \quad (1.18)$$

with particle masses:

$$M_W^2 = \frac{1}{4}g^2\nu^2 \quad (1.19)$$

$$M_Z^2 = \frac{1}{4}(g^2 + g'^2)\nu^2 \quad (1.20)$$

$$M_A = 0 \quad (1.21)$$

---

<sup>3</sup>Unitary gauge transformation is to select the fixed gauge which sets the Goldstone boson terms into 0

From the expression, it turns out that  $Z$  boson and photon are both the mix of  $B$  and  $W^3$  bosons with different phases which could be shown as:

$$\begin{bmatrix} A \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{bmatrix} \begin{bmatrix} B \\ W^3 \end{bmatrix} \quad (1.22)$$

With  $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$  and  $\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$ . Here,  $\theta_W$  is called the weak mixing angle or Weinberg angle. By this, the electroweak parameter,  $\rho$ , is defined:

$$\rho = \frac{m_W}{m_Z \cos\theta_W} \quad (1.23)$$

with the comparison between Eq. 1.2 and Eq. 1.14 with Eq. 1.22, the electric charge could be defined as:

$$e = g \sin\theta_W = g' \cos\theta_W \quad (1.24)$$

This relation gives the access to a precision measurement of  $\rho$ , which is now given 1.0008, a little deviation from expectation of 1 in the SM because of the loop diagram correction.

In terms of degrees of freedom, before symmetry breaking, its comes with four degrees from Higgs complex scalar doublet, six degrees from  $SU(2)_L$  gauge field,  $W_i$ , and two degrees from  $U(1)_Y$  gauge field,  $B$ , which makes 12 degrees in total for all the massless fields. After symmetry breaking, the number of degrees of freedom doesn't reduce with nine degrees from three massive vector boson,  $Z$  and  $W_{\pm}$ , two degrees from massless photon,  $A$ , and one degree from physical real scalar field,  $h$ .

Not only granting mass to bosons, the interaction between fermions and Higgs boson is also part of the Brout-Englert-Higgs Mechanism. The left-handed fermionic field is

defined as a doublet:

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad (1.25)$$

For right-handed leptons, the representation would be in a singlet because of the lack of right-handed neutrinos.

Their interaction with Higgs field are through Yukawa couplings<sup>4</sup>

$$\mathcal{L} = -\lambda \bar{Q}_L \Phi d_R + h.c. \quad (1.26)$$

with  $\lambda$  as the coupling constant. The Lagrangian can lead to the quark mass as:

$$m_d = \frac{\lambda v}{\sqrt{2}} \quad (1.27)$$

This mechanism would change the chirality of a fermion, when it is giving the mass. However, no right-handed neutrino and left-handed anti-neutrino were measured, which leaves it as one of the unsolved problem in SM. (More details are given in next section.)

### 1.3 Unsolved Problems in SM

With SM, we have understood most behaviours of the fundamental particles. However, it still failed explaining some experimental results. The following is part of the them the work in the thesis is trying to answer.

#### Higgs Mass Naturalness

---

<sup>4</sup>Yukawa coupling means the couplings between fermionic and bosonic fields

In quantum field theory, all the experimental observables could be presented as:

$$O = a_1 + a_2 + a_3 + \dots \quad (1.28)$$

where  $O$  corresponds to the physical observables like the invariant mass of particles, and  $a'_n$ s are the independent contributions to the observables. For naturalness of the observable, it is expected that  $a_n \leq O$ . For any case that  $a_n \gg 0$ , the further fine-tuning needs to be introduced for proper correction on theory, and it also indicates the defect in the theory.

The form for the observable of Higgs mass is:

$$m_h^2 = 2\mu^2 + \delta m_h^2 \quad (1.29)$$

where  $\delta m_h^2$  for the contribution from coupling to top quark is:

$$\delta m_h^2 \simeq \frac{3}{4\pi^2} \left( \lambda_t^2 + \frac{g^2}{4} + \frac{g^2}{8\cos^2\theta_w} + \lambda \right) \Lambda \quad (1.30)$$

where  $\lambda_t$  is the top-quark Yukawa coupling,  $g$  is the  $SU(2)$  gauge coupling,  $\lambda$  is the coupling constant in the quadratic term in Higgs potential and  $\lambda$  is the energy cut-off to divergent loop integrals. With the observed Higgs boson mass at  $125\text{GeV}$ ,  $\Lambda$  is estimated to be around  $1\text{TeV}$ , and that is also roughly the limit to keep the naturalness of this observable.

However, many models beyond the SM predicts the existence of particles at the TeV scale, which means the naturalness would be broken in the scenario. For this reason, a correction for Brout-Englert-Higgs Mechanism is needed, or there is possibly a heavier Higgs boson to complete the theory.

### The Hierarchy Problem and Quantum Gravity

The hierarchy problem is defined in two ways: the unreasonable discrepancy be-

tween theoretical prediction and experimental result, or two comparable parameters. Higgs mass is one instance for the first definition. For the second one, it is generally referred to the gap between coupling strengths of weak interaction and gravity for the order of  $10^{16}$ .

When a hierarchy problems occurs, the “so-called” fine-tuning is introduced to correct the discrepancy between two parameters. However, the fine-tinning could only be performed with enough understanding on the quantum effect of related parameters, and quantum gravity is still an unsolved problem. In the case, no solution is available.

### **Neutrino Mass**

Brout-Englert-Higgs Mechanism is the process to make particles massive within which the chirality of fermions would be changed. This implies that massive fermions of right-handed and left-handed chirality shall both exist, but no evidence is found for right handed neutrinos. Therefore, they are supposed be massless with SM. However, with the measurement of neutrino oscillation induced by the difference of neutrino mass and flavour eigenstates, they are practically massive particles. The conflict between SM and experiment still remains unsolved.

## **1.4 Thesis Overview**

To solve the problems in SM, analyses are performed in two ways, resonance and non-resonance searches which are corresponding to two different signatures in physics: new particles or new couplings. The thesis will present how the experiment is set up to see the signatures of new physics in Chapter 2, and the following 4 chapters are dedicated to show the analyses of resonance and non-resonance searches with 2015+2016 data. The last chapter is for the simulation of the upgrade of LHC and ATLAS detector which will start to operate in 2021.

# Chapter 2

## Experimental Setup

The accelerators are utilized to recreate the high energy environment rich in new physics like the hot early universe. In this thesis, Large Hadron Collider (LHC) is used for this purpose, and the ATLAS detector (A Toroidal LHC ApparatuS) is taken to probe the potential signatures of new physics.

### 2.1 Large Hadron Collider

LHC is a circular collider with diameter of 27km for hadrons (it could be either proton or lead ion) hosted by CERN at the border of France and Switzerland in the depth varied between 50m to 175m. It accelerates protons (lead ions) to the speed of Lorentz Factor of 10540 (32) and smashes them together to recreate the “hot” environment right after the big bang which corresponds to 6.5TeV (2.5TeV) energy. However, before a proton reaches the targeted energy, it has a long way to go.

#### Ionization

At the beginning, hydrogen is released from a tank and ionized into the state of proton-electron plasma. It then experiences the electric field to separate electrons as well as protons like Fig. 2.1. The protons are then taken out and sent into the LINAC2, a linear accelerator. After reaching the energy of 450MeV, the protons are fed into

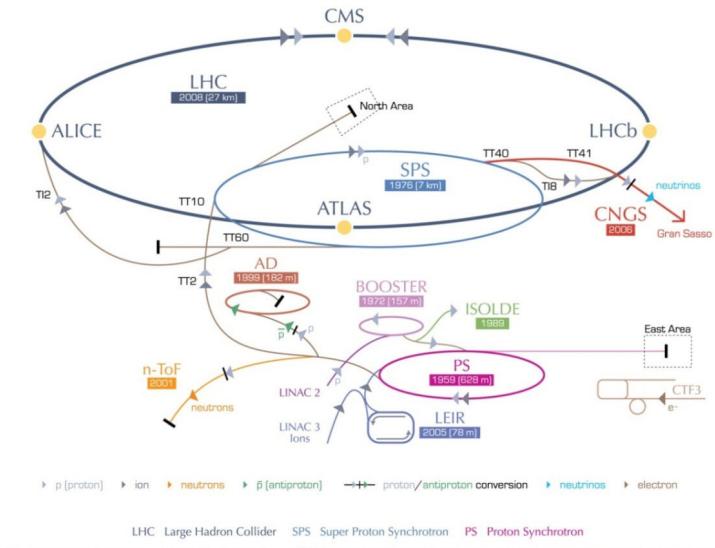
circular accelerators in the order of the PSB, PS and SPS to further increase the energy until they reach 450GeV (Fig. 2.2). By this stage, the protons are ready to be injected into the LHC.



**Figure 2.1:** The hydrogen plasma is separated into electrons (red) and protons (blue)

## Magnets

### CERN's accelerator complex



European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

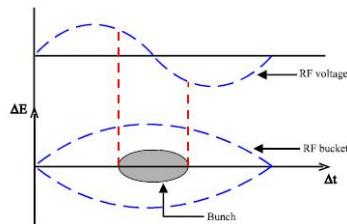
© CERN 2008

**Figure 2.2:** Before the LHC, protons go through several boosting facilities

While accelerating the protons, they would repel each other due to the same electric charge they are carrying, so the quadrupole magnets are implemented in LHC to focus them by the effect of magnetic lens. In addition to the quadrupole magnets, the other magnet system in LHC is the superconducting dipole magnets working to bend the protons to keep them staying in the circular pipe of the LHC. The dipole system was upgraded between 2012 and 2015 to provide a 8.3T magnetic field to bend the proton beam at an energy of 6.5TeV.

### Radiofrequency Cavity

The “radiofrequency cavity” (RF cavity) is in charge of the acceleration. Protons would experience electric field when going through RF cavities which are installed in the LHC like beads along a string. The field is induced by an alternating current of a frequency of 400 MHz and resonates as a standing wave in the cavity. This wave decelerates faster protons and accelerates slower ones, which makes the protons squeeze into bunches as demonstrated in Fig. 2.3., until they reach the targeted energy. When the beams are kept in the same speed, they are called “stable beams” and ready for the collision. Each LHC beam could have up to 3564 bunches with  $\mathcal{O}^{12}$  protons in each



**Figure 2.3:** The protons are formed into a bunch in the EM wave

bunch for a spacing of 25ns, but not all of them are filled. For the LHC 2018 operation, the “filling scheme” has around 1000-2500 bunches filled, while the remaining ones are left empty (filling scheme most of time is constrained due to technical issues). A series of continuous bunches is called a “bunch train”. This scheme would then be used to configure the trigger and data acquisition system for the active window of detector operation.

## Collision

The LHC has two beams going in opposite directions with the same configuration (bunch structure, luminosity and energy), and the two beams cross at locations where four detectors are sited: ALICE, ATLAS, CMS, and LHCb. Before stable beams, the two beams pass each other where they are supposed to cross. When both of the beams are ready, the two beams are slightly shifted to target on each other for the collisions . The crossing angle between the two beams plays an important role in detector performance. It should not be too big, or it would have an impact on physical object reconstruction (see section. 2.3) which assumed a zero crossing angle. However, it also should not be too small, or the two beams would interfere with each other. The crossing angle is kept optimized during LHC operation even when the detectors are taking data for physics.

When collisions happen, the two crossed bunches usually have more than one pair of interacting protons. In physics, only the most energetic one gets the attention for study, while the other ones are the background contribution called "pile-up events". For ATLAS 2018 operation, the pile-up events could number up to 70 per bunch crossing, and it is now a major challenge of analyses to suppress this type of background.

## 2.2 ATLAS Detector

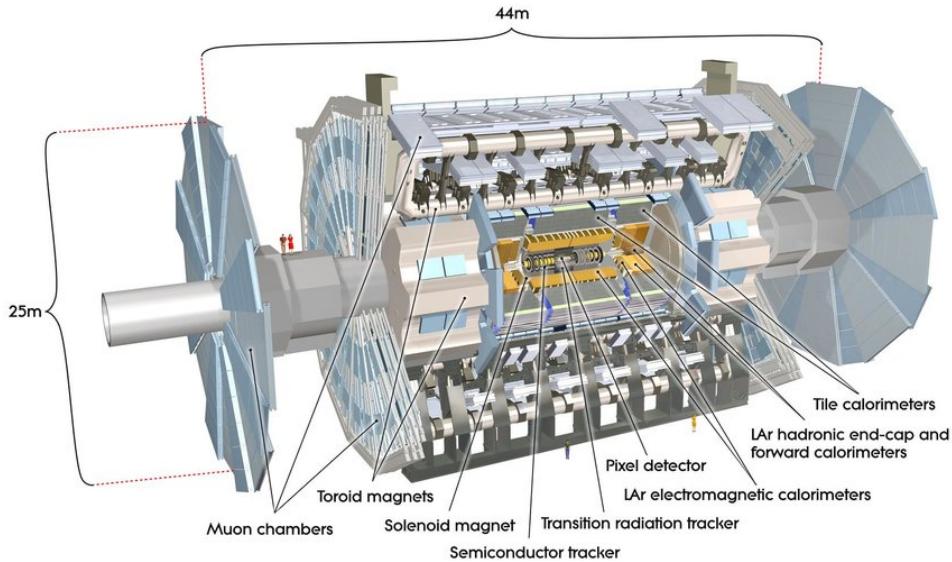
The ATLAS detector (A Toroidal LHC ApparatuS) is designed as a general purpose detector<sup>1</sup> aiming for most high energy physics topics in the energy scale LHC provides like SM precision measurement and searches for new physics.

The ATLAS detector is in a cylinder shape with dimensions of 44m in length and 25m in diameter. Its inner structure is like an onion with multiple layers from the inner most tracking system to the outer part of muon spectrometer functioning to capture different physical objects which will be explained in the following. In the purpose of measuring the particle mass and charge, ATLAS also has two magnetic systems (a

---

<sup>1</sup>The other general purpose detector hosted by LHC is Compact Muon Solenoid (CMS). The discovery of any new physics shall be verified by both the ATLAS and CMS collaborations

solenoid and a toroid) located outside inner tracking system and muon spectrometer. The diagram of the whole ATLAS detector is shown in Fig. 2.4.



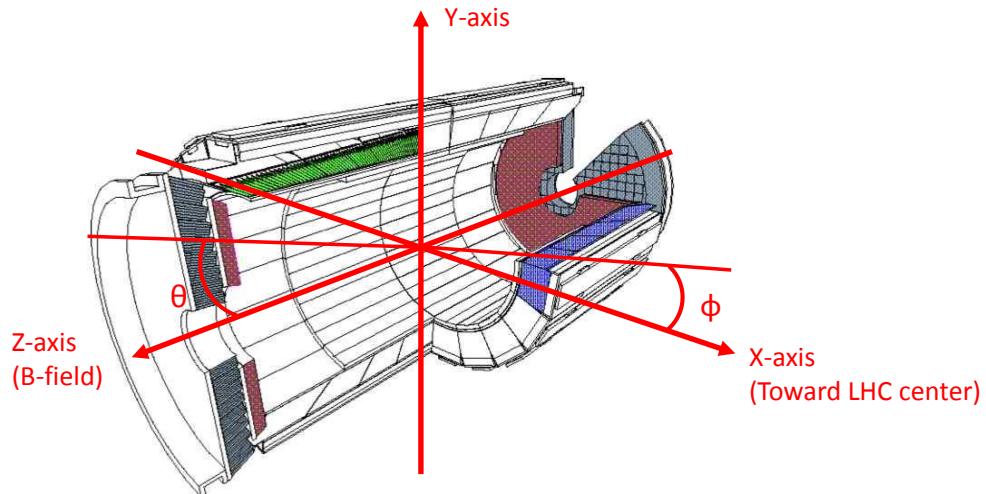
**Figure 2.4:** The diagram of the ATLAS detector

To define the object positions inside this massive and complicated giant, the coordinate is applied as shown in Fig. 2.5. The x-axis is defined pointing to the centre of the LHC, while the z-axis is the cylinder axle toward the direction of solenoid magnetic field. Then, the y-axis could be found with the right-hand rule. However, this Cartesian coordinate is not convenient in a cylinder, so, instead, the spherical system ( $\theta$ :angle related to z-axis,  $\phi$ :angle related to x-axis) is adopted in terms of physics. To keep the parameters as Lorentz invariance,  $\theta$  is interpreted into pseudorapidity,  $\eta$ :

$$\eta = -\ln \tan \frac{\theta}{2} \quad (2.1)$$

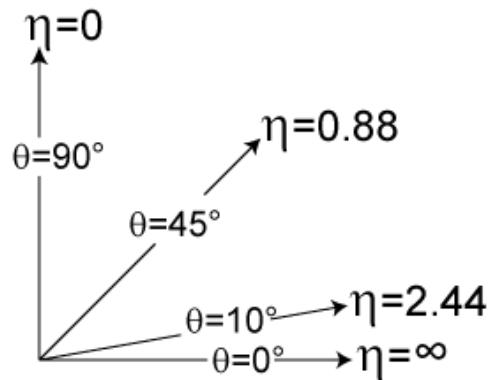
With this definition, the variation of  $\eta$  is different from  $\theta$ , which can be seen in Fig. 2.6. This quantity is important, because the distance between two particles in the detector is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2.2)$$



**Figure 2.5:** The coordinate system used in the ATLAS detector

For the same  $\Delta R$ , the separation would be actually larger in the high  $\eta$  region especially at  $|\eta| > 3.2$  (“endcap” and “forward” regions).



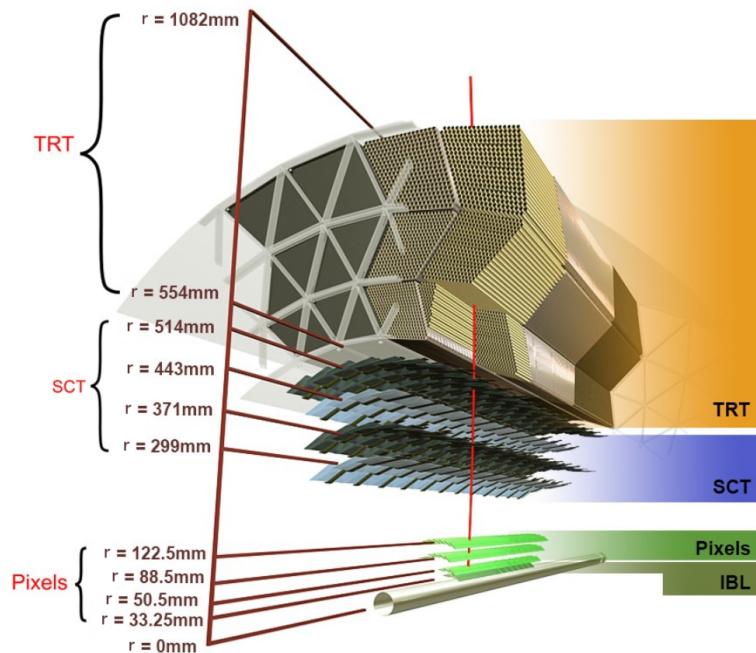
**Figure 2.6:** The Psedorapidity varied with  $\theta$

### 2.2.1 Inner Detector (ID)

The design of a general detector usually consists of two types of system: “trackers” and “calorimeters”. The tracker is used to record the particle trajectories inside the

detector with the lowest disturbance on its energy, while calorimeters trap the particles to measure its total energy sum,  $E$ .

The ATLAS Inner detector is designed as a “tracker”, so it is used to take the tracks of particles from the collisions. It stands at the inner most part of the detector and spans from 3cm to 108cm in radius with several layers from three subsystems which are pixel, semiconductor tracker (SCT) and transition radiation tracker (TRT) as shown in Fig. 2.7.



**Figure 2.7:** The diagram for the ATLAS inner detector

Each layer has cells of well-defined granularity. When particles are passing through the inner detector, they leave “a hit” per cell on each layer. The tracks are then defined as the link through hits on each layer which are curved lines due to the existence of magnetic field from solenoid, so the curvature of a track could be taken to evaluate the particle momentum and charge. After all tracks are reconstructed, the vertexes are then defined as where the tracks cross. The resolution of transverse momentum,  $p_T$ <sup>2</sup>,

<sup>2</sup>In ATLAS, the activity on transverse plane (i.e. x-y plan) has most of physics interest, because the transverse momentum sum is supposed to be 0, but the case for longitudinal direction isn't

depends on the particle  $p_T$  and  $\eta$ , and it can be presented as:

$$\sigma_{p_T} = \sqrt{a^2 p_T^4 + b^2 p_T^2} \quad (2.3)$$

with  $a$  and  $b$ , the coefficients, depending on track quality and  $\eta$ . From MC simulation for the track with least seven hits (a track crossing all layers from pixel and SCT) within  $0.25 < |\eta| < 0.5$ ,  $a$  and  $b$  are estimated to be  $0.00034 \text{ GeV}^{-1}$  and  $0.0015$  respectively.

## Pixel

The pixel detector is the innermost system of ATLAS, and it has the structure of three concentric barrels enclosed by three disks at each end, so all the particles coming out from the collision must pass through all the layers (giving three hits). It provides the best position resolution in the ATLAS detector with a granularity of  $50\mu\text{m} \times 400\mu\text{m}$  for each cell in the  $r\Delta\phi \times z$  plane with the coverage of  $|\eta| < 2.5$  which is used to define the barrel region which has a spatial resolution of  $14\mu\text{m} \times 115\mu\text{m}$

In 2014, a new layer of pixel detector called insertable b-layer (IBL) was installed at  $3.3\text{cm}$  to the beam pipe in addition to the original three layers. Its design is aiming to assist with measurements of short-live particles (like the b quarks whose lifetime is  $10^{-12}\text{s}$ ), so it has an even better granularity of  $50\mu\text{m} \times 250\mu\text{m}$  with extended coverage to  $|\eta| < 3$ . The improved granularity also helps to reduce the uncertainty on impact parameter of collisions.

## Semiconductor Tracker

Outside the pixel detector is the semiconductor tracker with three layers in its barrel and nine disks at each end. The sensors are double sided, so when a particle passes through four layers, it leaves totally eight hits in the SCT which form four spacepoints. Different from the pixel detector which has one sensor on each module, the SCT modules have two strip sensors with the width of  $80\mu\text{m}$  which cross at an angle of  $40\text{mrad}$  giving a spatial resolution of  $17\mu\text{m} \times 580\mu\text{m}$  in the  $r\Delta\phi \times \Delta z$  plane.

## Transition Radiation Tracker

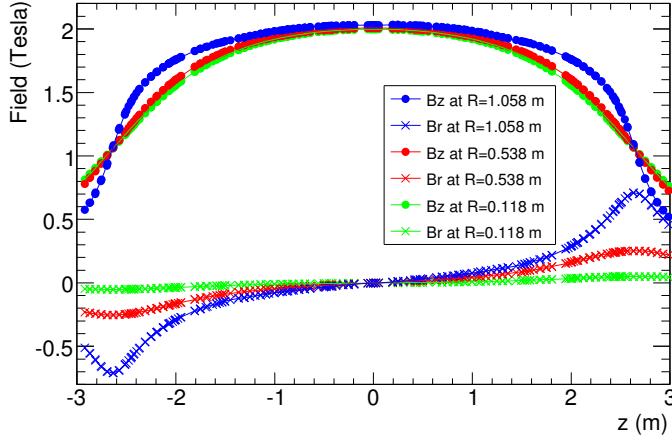
The last part of the inner detector is the TRT detector. It doesn't have a multiple layer structure as the pixel or SCT detectors but just a single thick layer stacked of straw drift tubes. Each straw has the diameter of 4mm (with the drift time correction, the spatial resolution from each measurement is  $130\mu m$ ) and is filled with the gas mixture of  $Xe$ ,  $CO_2$  and  $O_2$ . The gas mixture is used to optimize the absorption of transition radiation. (Due to the gas leaking problem found in the ATLAS operation from 2009 to 2012, part of the gas was replaced by cheaper  $Ar$ -based gas.) When a charge particle passes through the gas, the emitted photon (transition radiation) induces a "charge avalanche". This detector allows to distinguish between electrons and charged pions (because light particles emit more transition radiation).

## Magnets

The ATLAS detector has two superconducting magnet systems different from the CMS experiment with only one solenoid magnet. The inner one is the solenoid magnet located between the TRT detector and the calorimeter, while the toroid magnet is situated in the muon spectrometer system. The advantage of this design is to have the light material (solenoid) inside the detector for transparency, and the toroid still provides the magnetic field to further improve the resolution of momentum measurement.

The solenoid magnet is with a diameter of  $2.56m$  and length of  $5.8m$ . The magnetic field inside the solenoid is almost uniform of  $2T$  along the z-axis as shown in Fig. 2.8 to give the momentum and charge measurement in the inner detector.

The toroid magnet is composed of the barrel and endcap toroids, and both of them have eight coils providing the magnetic field of  $4T$  in the muon spectrometer. The toroid magnet has the advantage that the particle trajectories in the transverse plane are always perpendicular to the magnetic field, so the momentum measurement is



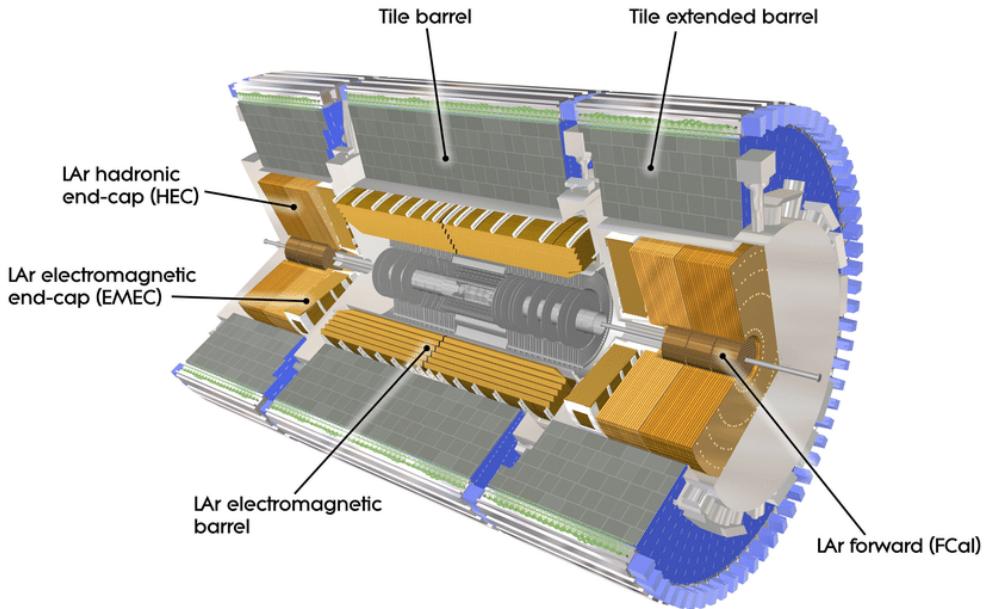
**Figure 2.8:** The magnetic field inside the solenoid

simplified. The toroid magnet is for the measurement of muon momentum in the muon spectrometer.

## 2.2.2 Calorimeter

Outside the inner detector is the calorimeter, an energy sampling system. In the ATLAS analyses, there is the need to distinguish the fundamental particles with their energy, so two systems of calorimeters are applied to trap particles with different mass: the electromagnetic calorimeter (ECAL) for electrons and photons as well as the hadronic calorimeter (HCAL) for the hadronic particles. Both ECAL and HCAL have the coverage up to  $|\eta| < 4.9$ . For the range of  $|\eta| < 2.5$  in the barrel region, two types of calorimeter, the liquid argon (LAr) and tile detectors, are used for the ECAL and HCAL, while in the endcap and forward regions is only the LAr detector. To fit into the cylinder shape of the ATLAS detector, the calorimeters are accordion-shaped from the cross section side. The full diagram of the calorimeter system is presented in Fig. 2.9. The energy resolution for the calorimeter could be presented as:

$$\sigma(E) = \sqrt{a^2 + b^2 E + c^2 E^2} \quad (2.4)$$



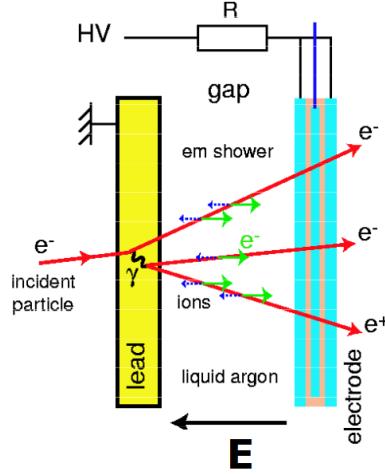
**Figure 2.9:** The calorimeter system of the ATLAS detector

where  $a$ ,  $b$ , and  $c$  are the coefficients. The first term is due to the electronic noise (constant), and the second term is from the shower development of the Poisson fluctuation for the number of shower particles, while the third terms is for the calorimeter non-uniformities (linear to the true shower energy). From the test beam data, the coefficients for the ECAL are  $0.4\text{GeV}$ ,  $0.1\sqrt{\text{GeV}}$  and  $0.0017$  for  $a$ ,  $b$ , and  $c$ . In terms of the HCAL, the resolution is a bit worse with  $1.6\text{GeV}$ ,  $0.52\sqrt{\text{GeV}}$  and  $0.03$ . The degraded resolution is due to the intrinsic property of the measurement on hadronic objects which have the energy contribution from neutrinos or binding energy between hadrons.

### Electromagnetic Calorimeter

In ATLAS, the ECAL is made up of the LAr detector each module of which has one absorber and one electrode, and liquid argon is the medium between them. When a particle hits the absorber, it induces the shower, and the shower electrons ionize liquid argon atoms. All the electrons from the interactions would then be collected by the electrodes. The measured current is used to estimate the energy of the incoming

particle. The process could be seen in Fig. 2.10.



**Figure 2.10:** The interaction between an electron and the LAr calorimeter

The barrel LAr detector has three sampling layers with different depth and granularity. An extra presampler (layer 0) was added for  $|\eta| < 1.8$  which has no absorber but only a thin LAr sampler to recognize photons from  $\pi^0$  decays. The best granularity is at the strip layer (layer 1) for  $0.0031 \times 0.1 (\Delta\eta \times \Delta\phi)$ <sup>3</sup>, while the last layer is coarse for  $0.05 \times 0.025$  in terms of  $\Delta\eta \times \Delta\phi$ . For the energy absorption, the depth is what matters most. The full depth of the three sampling layers could correspond to  $\sim 22$  lead radiation lengths ( $X_0$ ) or 2 nuclear interaction length ( $2\lambda$ ).<sup>4</sup> When the ECAL is extended to the region of  $2.5 < |\eta| < 3.2$ , only the last two layers would remain, but they still have  $18X_0$  in total.

### Hadronic Calorimeter

Behind the LAr detector is the three-layer tile detector covering  $|\eta| < 1.7$  with a crack<sup>5</sup> at  $1.37 < |\eta| < 1.52$ . It operates in the similar way to the LAr detector, but the absorber material is scintillator. Each sensor of this system is coarser as compared

<sup>3</sup>the granularity for  $\Delta\phi$  is a approximation, as it has to complete a circle of an irrational number.

<sup>4</sup>the radiation length is defined by the electron energy loss, while the nuclear interaction length is defined by the hadronic object energy loss.

<sup>5</sup>The crack is for the supporting structure and output cables

to LAr ones with  $0.1 \times 0.1$  ( $\Delta\eta \times \Delta\phi$ ) for the first two layer and  $0.1 \times 0.2$  at the third layer. With the need for the absorption of hadronic objects, it has  $8\lambda$  in the depth for all three layers.

In the endcap region ( $1.7 < |\eta| < 3.1$ ), another type of LAr detector with copper absorber is used as the HCAL. It contains four layers which have the same granularity for  $0.1 \times 0.1$  ( $\Delta\eta \times \Delta\phi$ ) in the region,  $1.7 < |\eta| < 2.5$ , and  $0.2 \times 0.2$  in  $2.5 < |\eta| < 3.1$ .

### Forward Calorimeter

In contrast to the inner detector, the calorimeter should capture as many particles as possible, so the missing energy carried by invisible particles could be estimated by energy conservation within the detector. Therefore, a forward detector is installed at  $3.1 < |\eta| < 4.9$ , and it has the best rapidity coverage among the ATLAS subsystems.

The type of detector used here is the third type of LAr detector with tungsten absorber. It has three layers with the first one for ECAL and the last two for HCAL with the same depth of  $10\lambda$  (ECAL+HCAL).

### 2.2.3 Muon Spectrometer

The outermost detector is the muon spectrometer (MS). Because of their large mass and lack of strong interactions, only muons could travel through the calorimeter and leave signatures here. The muon spectrometer is composed of four types of detectors: thin gap chamber (TGC), resistive plate chamber (RPC), monitored drift tubes (MDT), and cathode strip chamber (CSC) with the toroid magnet system.

In this subsystem, the MDT and CSC are the two detectors providing the tracking measurement with a three-layer structure. In the coverage of  $|\eta| < 2.0$ , all the three layers are composed of the MDT detector, while the innermost layer is replaced by the CSC detector in the extent of  $2.0 < |\eta| < 2.7$  for the effectiveness of high particle density environment. The overall tracking measurement has the spatial resolution of  $35\mu m$ .

However, the precise tracking measurement of the MDT and CSC comes with the cost of a poor temporal resolution, so the RPC ( $|\eta| < 1.05$ ) and TGC ( $1.05 < |\eta| < 2.7$ ) are interspersed in tracking layers with the time resolution of  $25\text{ns}$  (with the consideration of uncertainty from cosmic muons). With the fast response, they are part of the ATLAS hardware trigger system. The overall detector performance is summarised in Tab. 2.1.

**Table 2.1:** Muon Spectrometer Subdetector Performance

Type	Function	coverage	z/R resolution	$r\Delta\phi$ resolution	time resolution
MDT	tracking	$ \eta  < 2.7$	$35\mu\text{m}(z)$	N/A	N/A
CSC	tracking	$2.0 <  \eta  < 2.7$	$40\mu\text{m}(R)$	$5\text{mm}$	$7\text{ns}$
RPC	trigger	$ \eta  < 1.05$	$10\text{mm}(z)$	$10\text{mm}$	$1.5\text{ns}$
TGC	trigger	$1.05 <  \eta  < 2.7$	$2 - 6\text{mm}(R)$	$3 - 7\text{mm}$	$4\text{ns}$

## 2.2.4 Trigger System

The LHC has the collision rate at  $40\text{MHz}$ , which leads to the data rate over  $60\text{TB}$  per second. However, most of the events have no physical interest, because they are just the products of low energy hadronic interactions. Therefore, the trigger system is developed to select events which are going to the storage.

For data-taking, the ATLAS trigger system has a two-level structure: the hardware-based L1 trigger (L1) and the software-based high level trigger (HLT). The L1 system is based on the front-end electronics with the logic of selection written by FGPA. Its feature is to make a fast reconstruction of physical objects with a degraded resolution, and it delivers the events at the rate of  $100\text{kHz}$  ( $100k$  events per second). As the detector signatures from the calorimeter and MS are irrelevant to each other, they have their independent L1 trigger systems: L1Calo and L1MU. After the physical objects objects are reconstructed in the two systems, they are then sent to the L1Topo system for the estimation on the topological relation between them. The final trigger decision would eventually be made at the central trigger processor (CTP) by whether a

event contains the objects with energy or topological parameters fulfilling the defined criteria. Afterwards, those objects are further processed with a more complicated reconstruction algorithm to give the HLT trigger decision, and the output event rate is reduced to  $\sim 1\text{kHz}$ . The full trigger system is shown in Fig. 2.11.

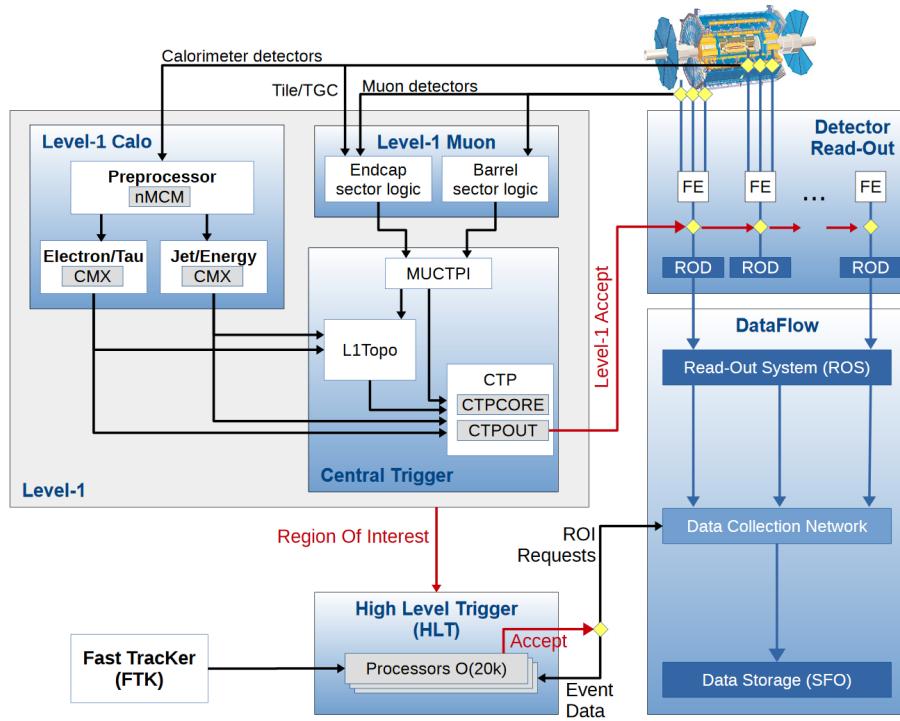


Figure 2.11: The ATLAS trigger system

## L1Calo

When the detector signatures are received from the calorimeter, they are firstly sent to the readout buffer and the L1Calo system. The first component in the L1Calo electronic system is the preprocessor where the signatures are processed into trigger towers with degraded granularity and sent to the processors for physical object reconstruction. Electrons, photons and taus are reconstructed with the trigger tower of  $0.1 \times 0.1 (\Delta\eta \times \Delta\phi)$  in the cluster processor (CP), while hadronic objects and missing transverse energy ( $E_T^{\text{missing}}$ )<sup>6</sup> are processed in jet energy processor (JEP) with a coarser

<sup>6</sup>As the protons only have the longitudinal momentum, the transverse direction momentum should be conserved after collisions

granularity of  $0.2 \times 0.2$ .

## L1MU

The L1MU system is taking the data from the RPC and CSC which have great time resolution as fast as  $1.5\mu s$  but with a poor spatial resolution. It receives signatures from the MS barrel and endcaps where they are processed respectively. To further suppress the rate contributed by fake muons, the L1 muons are reconstructed with consistent hits from the TGC at the endcap ( $1.05 < |\eta| < 2.7$ ).

## HLT

When the trigger decision was made to accept an event, the regions of interest (ROI) with the original detector granularity are passed to the HLT. The HLT is runs on a CPU farm where the more complicated algorithms are deployed to reconstruct the physical objects. Due to the finer granularity and longer latency, it provides better precision on both energy and spatial resolution. When the events fulfil the HLT criteria, they are then sent to storage.

## ATLAS Trigger Menu

An ATLAS trigger is generally a trigger chain composed of L1 and HLT items. When an HLT trigger is fired, there is always a corresponding L1 trigger decision. For example, HLT electron trigger shall only be passed when a L1 electron trigger is also fired:

$$L1\_e24 \rightarrow HLT\_e26\_lhtight\_nod0\_ivarloose \quad (2.5)$$

where the numbers are the trigger thresholds in the unit of  $GeV$ , while the *lhtight* and *ivarloose* are to define the electron quality with the calorimeter activities in the surrounding region of this electron (see more details in Sec. 2.3). The threshold of triggers might not be kept the same during the operation at periods. Because the LHC keeps pushing its performance on instantaneous luminosity, the energy contribution

from pile-up events enhances the trigger rate above the allowed bandwidth for data storage. To make better suppression on the trigger rate, the thresholds are therefore raised during some operation periods.

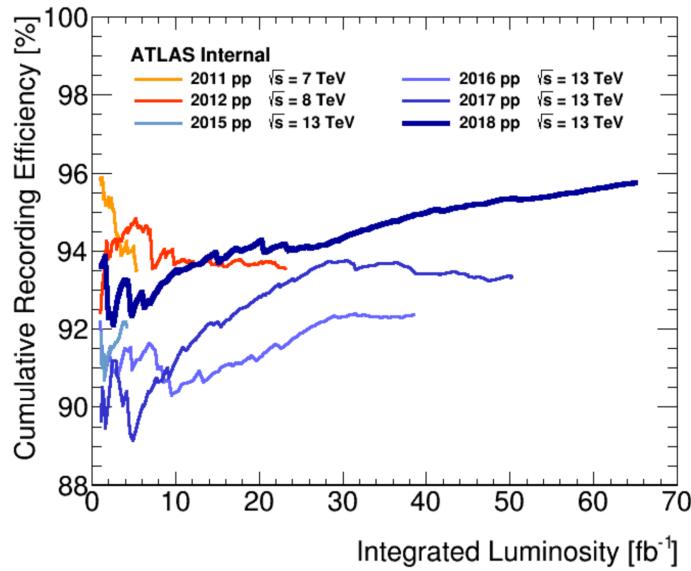
The defined triggers would then be made into “streams” where the events are categorized for different purposes. Physics analyses shall use the triggers contained in the “*physics\_main*” stream, and there are also the dedicated streams composed of “prescaled” triggers for hardware calibrations. Those calibration triggers usually have lower thresholds in contract with the ones in *physics\_main*. A random sampling is applied to only pick a fraction of events passing those triggers, which makes them unsuitable for physics analysis. The total allowed output rate from all streams is 5kHz with 1kHz for *physics\_main*.

### 2.2.5 Run 2 Operation Overview

For the ATLAS operation from 2015 to 2018 which is called Run 2 (with respect to Run 1 from 2009 to 2013), the average data recording efficiency is around 95% with respect to LHC delivery efficiency. The integrated luminosity is  $\sim 36\text{fb}^{-1}$  for 2015 and 2016,  $\sim 46\text{ fb}^{-1}$  for 2017 and  $\sim 63\text{ fb}^{-1}$  for 2018 which gives the total data of  $140\text{fb}^{-1}$ . The performance from 2011 to 2018 is summarized in Fig. 2.12

## 2.3 Object Reconstruction

When the events are passed into the permanent storage, they are still in the format of raw data which contains only the information of hits (the spacepoints from the inner detector and muon spectrometer) and calorimeter energy towers (the energy deposited in the calorimeter cells). They need to go though the full reconstruction (offline reconstruction) to be interpreted into the objects with physical meanings as the SM particles like electrons or muons. The reconstruction will be based on the principles of interactions between detectors and particles:



**Figure 2.12:** The ATLAS recording efficiency and luminosity

- 1) Only charged particles leave tracks in the inner detector
- 2) The charged particles shall deposit energy in the ECAL, and the light ones are stopped here.
- 3) All the particles except for muons are supposed to be stopped in or before the HCAL.
- 4) Only muons could reach the muon spectrometer.

After the reconstruction of all subjects, a further correction on energy scale (the peak of the energy pulse shape) is applied on both data and simulation samples to take in the effect of energy loss from the radiation, the contamination from other objects, or the detector effect (like dark current, hot noise, or material inhomogeneities). The final procedure is to remove the overlapped objects by the priority defined by the analyses.

### Primary Vertex & Tracks

A pattern recognition is performed in the SCT to find the helical trajectories with at least 3 spacepoints and  $p_T > 500\text{MeV}$  which are taken as the track seeds. A Kalman Filter algorithm is then performed to extend the track seeds to the pixel layers. To resolve the reconstruction ambiguity, a tracking score system is taken to reject the shared spacepoints or fake tracks. When a track is reconstructed with more hits and less “holes” (missing hits in some layers), it is given a higher score. The final tracks shall all have at least seven hits (three spacepoints from SCT and four hits from pixel). To complete the track reconstruction, the projection route from the outermost SCT spacepoint of a track passes through TRT, and the drift tubes within  $10\text{mm}$  from the route are integrated into the track. Afterwards, one addition track reconstruction (outside-in) from TRT is performed to recover the tracks from the late decay or photon conversion. The unused TRT segments are then rematched to the SCT and pixel hit remnant.

The crossing points of tracks near the LHC pipelines are then assumed to be where the interactions happen, which are called “vertices”. The vertex associated with the highest  $p_T^2$  sum is then defined as the primary vertex. All the reconstructed objects should origin from the primary vertex which is verified by  $d_0$  and  $d_z$ , the distances between the object and the primary vertex in the x-y plane and the z-axis, or they are taken as “minimum bias” background.

## Electrons

Electrons are charged light particles, so they leave tracks in the inner detector and energy clusters in the ECAL. The two types of signature are combined and testified to reconstruct electrons.

The first stage of the reconstruction is to build the energy cluster as an electron seed. A window of  $3 \times 5$  ECAL layer-2 cells (corresponding to  $0.075 \times 0.125$  in  $\Delta\eta \times \Delta\phi$ ) is used to scans through ECAL layer-2 to find the electron seeds. If the transverse energy sum ( $E_T$ ) inside the window is above  $2.5\text{GeV}$ , the cells inside this window are selected, and the electron position is defined as the energy weighted  $\eta$  and  $\phi$

(barycentre) of this window. Then, this window is extended along the  $R$ -direction to sum over the energy in other layers with the adjusted window size (detailed in [1]). The cells taken in the cluster are then removed to avoid the duplication into other electrons. With the estimation from  $Z \rightarrow ee$  simulation sample, this algorithm has the electron reconstruction efficiency of 95% ( $> 99\%$ ) for  $E_T \sim 7\text{GeV}$  ( $E_T > 15\text{GeV}$ ).

The reconstruction of tracks associated to electrons is performed independently from the mentioned reconstruction. A track seed of  $p_T > 1\text{GeV}$  is firstly reconstructed with three spacepoints from the SCT layers. Then, based on the pion hypothesis (pion energy loss pattern in the ID materials), it is verified by whether this track seed can be extended to pixel with four hits and matched to a calorimeter cluster. If it fails, the electron hypothesis is applied for the same verification. The hits from both hypotheses are then fitted using “ATLAS Global  $\chi^2$  Track Fitter” into tracks, and the tracks failing the pion track hypothesis are then tested again with electron hypothesis. The tracks passing the electron hypothesis are then taken as potential electron tracks. This algorithm is also integrated in the standard track reconstruction with the least interference.

The track and cluster are then associated with a loose  $\Delta R$  matching which considers the electron bremsstrahlung and the number of hits in the inner detector. The matched track-cluster pairs are then refitted with optimised “Gaussian Sum Filter” (GSF) to take non-linear bremsstrahlung into account.

To have further separation between signal-like and background-like electrons, the electron identification is then performed on  $Z \rightarrow ee$  (signal) and dijet (background) MC samples. It is a multi-variable analysis (MVA) based on the likelihood discriminant defined as:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B} \quad \text{and} \quad \mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^n P_{S(B)}^i(x^i) \quad (2.6)$$

where  $P_{S(b)}^i(x^i)$  is the probability density function for a specific input variable  $x^i$ , and  $\vec{x}$  is the vector formed by them in the likelihood phase space of all input variables (all the input variables could be found in [ ]). Three working points are therefore defined by  $d_{\mathcal{L}}$ : *Tight*, *Medium* and *Loose*<sup>7</sup>. The signal efficiency from this selection process is as a function of electron  $E_T$ , and the plateau of efficiency could be reached at  $E_T$  70GeV for 97% (95%) [91%] on *Loose* (*Medium*) [*Tight*] working point.

In addition to the reconstruction quality, the electrons are also required to be “isolated” from all the other tracker and calorimeter signatures, because of the concern that the nearby detector activities might affect the electron measurement. The isolation is defined in two ways:

- the calorimeter isolation ( $Iso^{E_T}$ ): it is defined as the cluster  $E_T$  sum within a cone with  $R = 0.2(0.3)$  centred at the reconstructed electron inside which a central cluster subset in a rectangle of  $0.125 \times 0.175$  ( $\Delta\eta \times \Delta\phi$ ) is subtracted.
- the track isolation ( $Iso^{p_T}$ ): it is defined as the  $p_T$  sum of tracks from primary vertex within a cone of  $R = \min(0.2(0.3), 10\text{GeV}/E_T^e)$  centred at the electron but without the electron associated tracks.

The isolation discriminant is then applied as  $Iso^{E_T}/E_T$  or  $Iso^{p_T}/p_T$ . The recommendation working points on the discriminant are given as a function of  $E_T^e$  or fixed cut which are summarised with the muon isolation working points in Tab. 2.2.

## Muons

Muons are heavy enough to travel through the calorimeter and reach the MS, but the reconstruction is mainly based on the tracks in the inner detector and the MS.

The MS track segments are firstly built from the hits within each MS module, but the reconstruction coordinates are different in each subsystem: the MDT reconstruction is on the coordinate of the toroid magnetic bending plane, while the RPC and TCG have the coordinate orthogonal to it, and the CSC is only using the detector  $\eta$ - $\phi$  coordinate.

---

<sup>7</sup>*Tight* selected electrons are the subset of *Medium*, and *Medium* is the subset of *Loose*

A loose criteria is applied in the segment building algorithm to verify the compatibility to a full track. Then, the segments in the middle layer of the MS are taken as the track seed and extended to the inner and outer layers. If two segments could be fitted with enough hits by matching from their relative position and angle, they are integrated into the same track. The exception is in the transition region between the barrel and endcap, and a standalone and good-quality segment could be kept as a single track.

An overlap removal is afterwards applied to remove the shared hits in the tracks with poor fitting quality, but they could still be kept only if the fitting criterion is fulfilled. Two tracks could share maximally two hits in the inner two layers and have no same hit in the outer layer for the concern of close-by muons.

The hits along the tracks are then taken into the global  $\chi^2$  fitting. The hits with great deviation from the fitted MS trajectory are removed, and the fit is applied again to derive the new track. If there are hits not included in the track but within the allowed deviation from the track, they are also taken into the track, and the fitting is repeated.

The final MS tracks are taken as the seed to match to the inner detector tracks to reconstruct the combined muons. A further global fitting is conducted to extrapolate the muons with the flexibility to add in or remove the MS hits to improve the fitting quality with the ID tracks. The primary algorithm in the fitting is performed outside-in from the MS to the inner detector, and a complementary algorithm of inside-out is also applied to guarantee the robustness of the reconstruction. For the muons outside of the inner detector coverage ( $2.5 < |\eta| < 2.7$ ), they can be reconstructed from only MS tracks, but the criteria are more stringent.

Similar to electrons, muons also have the identification procedure with three parameters:  $q/p$  significance (the ratio of charge and momentum measured in the ID and MS over the quadrature sum of their uncertainty),  $\rho'$  (the ratio of momentum difference between the ID and MS measurements over the combined measurement) and the

normalised combined track fit,  $\chi^2$ . The working points for the muons identification have the definition individually as below:

- *Medium* muons: they are defined within the range of  $0.1 < |\eta| < 2.5$  with at least two layers of  $\geq 3$  hits. If it is within the range,  $0.1 < |\eta|$ , it is allowed to have hits in only one layer, but there shall be no hole in the MS track reconstruction. As the muons go beyond the coverage of the inner detector (i.e.  $2.5 < |\eta| < 2.7$ ), they shall have the MS tracks reconstructed from all three layers. An extra requirement of  $q/g$  significance above seven is also applied on this muon quality.
- *Loose* muons: those muons are defined with the most loose requirement. They are generally *Medium* muons, but the selection is loosen for the range of  $|\eta| < 0.1$  due to the missing coverage of the MS (where a gap is present for the service of the ID and calorimeter). When an ID track is found within this range and matched to a calorimeter cluster which is identified as a deposit by “minimum-ionization” particles, they are also accepted as loose muons to recover the reconstruction efficiency.
- *Tight* muons: all of them must have the tracks reconstructed from two layers in the MS (either MDT or CSC) with *Medium* muon hit selection. To enhance the purity of muons, a further requirement on the ID to MS track fitting is also added into the selection for  $\chi^2 < 8$ . An addition two-dimension cut on  $q/g$  significance and  $\rho'$  is also applied to improve the background rejection for muons with  $p_T < 20\text{GeV}$ .

For the muon isolation, the definition is similar to the electron ones, but they have different working points. The recommended working points for electrons and muons are shown in tab. 2.2.

## Jets

When quarks or gluons are travelling in the space, they went through the process called “fragmentation” or “hadronization” for “color-confinement” of QCD. This leads to the multiplication of quarks, gluons (i.e. partons) or even leptons and photons, and

**Table 2.2:** Electron/Muon Isolation Working Points

Working Point	Object	Calo Iso	Track Iso	Combined Iso
LoseeTrackOnly	all leptons	-	99%	99%
Loose	all leptons		99%	99%
Gradient	all leptons	$\epsilon = (0.1143 * p_T[GeV] + 92.14)\%$		$\epsilon(25GeV) = 90\%$ $\epsilon(60GeV) = 99\%$
GradientLoose	all leptons	$\epsilon = (0.057 * p_T[GeV] + 95.57)\%$		$\epsilon(25GeV) = 95\%$ $\epsilon(60GeV) = 99\%$
FixedCutTight	Electrons	topoetcone20/pT<0.06	ptvarcone20/pT<0.06	-
FixedCutTight	Muons	topoetcone20/pT<0.06	ptvarcone30/pT<0.06	-
FixedCutTightTrackOnly	Electrons	-	ptvarcone20/pT<0.06	-
FixedCutTightTrackOnly	Muons	-	ptvarcone30/pT<0.06	-
FixedCutLoose	Electrons	topoetcone20/pT < 0.2	ptvarcone20/pT < 0.15	-
FixedCutLoose	Muons	topoetcone20/pT < 0.3	ptvarcone30/pT < 0.15	-
FixedCutHighPtCaloOnly	Electrons	topoetcone20 < 3.5 GeV	-	-
FixedCutHighPtTrackOnly	Muons	-	ptcone20 < 1.25 GeV	-

they eventually form the bound states as hadrons leaving complicated signatures in the detector. One quark from a collision might leave more than one hundred tracks in the inner detector and several clusters in the calorimeter, and jets are defined as the ensemble of those signatures. To properly collect those tracks and calorimeter clusters into the same jets, the reconstruction algorithm is designed to ensure the infrared safety and collinear safety. The infrared safety means that the soft radiation from hadronic objects in a jet would not change the jet configuration like its width or the orientation, while the collinear safety indicates that the radiated objects which are not in the collinear direction of the jets would just be discarded to avoid an irrational jet width. To achieve both of the two requirements, the jet reconstruction algorithm employs the following two parameters for the jet definition:

$$d_{ij} = \min((p_t^i)^a, (p_t^j)^a) \times \frac{R_{ij}^2}{R} \quad \text{with} \quad d_{iB} = (p_t^i)^a \quad (2.7)$$

where  $p_t^i$  and  $p_t^j$  are  $p_T$  of  $i$ th and  $j$ th entities which could be calorimeter clusters (a patch of energetic cells in the calorimeter), tracks, or the truth particles from the simulation,  $R_{ij}$  is the distance between them, and  $R$  is the parameter to customize the algorithm for performance (i.e. cone size), while  $\alpha$  corresponds to three algorithms which are sensitive to different jet properties.  $\alpha = -2$  is for  $anti-k_t$  algorithm, and it has the advantage for better stability of jet structure during reconstruction with high sensitivity to hard objects and ignorance for the jet substructure as well as pile-up events.  $\alpha = 0$  and  $\alpha = -2$  are used for Cambridge-Aachen algorithm and  $k_t$  algorithm respectively which are more sensitive to jet substructure but with high dependence on pile-up events and soft objects. When  $d_{ij} < d_{iB}$ , the  $i$ th and  $j$ th objects are merged into the same cluster with the position defined as their barycentre. If no new pair could be found meeting this condition, the cluster is then defined as a jet. To make further

For the jets in the ATLAS experiment,  $anti-k_t$  algorithm is preferred with  $R = 0.4$  and  $R = 1.0$  with the inputs objects.  $R = 1.0$  is for the scenario that two jets are close to each other, and  $R = 0.4$  could not have the separation power to distinguish each of them. The input entities for jet reconstruction is the ECAL topo clusters in the range of  $|\eta| < 4.9$ , coverage of calorimeter, with energy above  $2\sigma$  with  $\sigma$  defined as the quadrature sum of pile-up events and electronic noise, while the seeding cluster to initiate the clustering has higher requirement of  $4\sigma$ . The hadronic energy deposit from HCAL is added to the ECAL clusters through “local cell weighting” (LCW). The final reconstructed jets are then used in analyses after the selection of “jet vertex tagger” (JVT) to guarantee that they originate from the primary vertex.

### b-Tagging

b quarks take over the greatest branch ratio of Higgs boson decay, and it is always the decay product of top quarks( $BR(t \rightarrow bW)=100\%$ ) which is the major background for most ATLAS analyses. Therefore, how to recognize the jets from b quarks is an important task for physics purposes.

Different from the signatures other stable SM particles leave in the detector, b-quarks

have a “pseudo-short” lifetime, so the jets of b-quarks have a displaced vertex several  $mm$  away from the primary vertex. The b-jet identification depends on this property and its decay chain which are evaluated in the following three algorithms:

- **Impact Parameter based Algorithms:** with the displaced vertex, the b-jet associated tracks would have a greater impact parameter in both transverse and longitudinal directions. They are taken into the probability density function for the ratio of possibility of b-quark and light-flavour quark hypotheses, which are then combined into a single log likelihood ratio discriminant (LLR).
- **Secondary Vertex Finding Algorithm:** to reconstruct the secondary vertex, the first step is to find the vertices with only two tracks. The tracks from the vertices far from the primary vertex, they are rejected. They might be from long-lived hadron decay (like kaon), photon conversion or hadronic interaction with materials. The secondary vertex is then reconstructed from the survived tracks with outlier tracks removed
- **Decay Chain Multi-Vertex Algorithm:** this algorithm is also called “jet finder”. Its purpose is to find the full chain of “ $PV \rightarrow b \rightarrow c$ ”. A Kalman filter is applied to link the vertices to approximate the trajectories of this jet with which the vertices of b- and c-quarks could be resolved even with only one-track link.

The output of those algorithms are then given to a multi-variable analysis (MVA) of Boosted Decision Tree for the final discriminant. It is trained with the sample of  $t\bar{t}$  which contains b-jets and the background composed of 10% c-jets and 90% other light-flavour jets. The final outcome, BDT score, is then applied as a simple cut to select b-jets, and the suggested working points are shown in tab. 2.3

### Missing Transverse Energy (MET)

The design of the ATLAS detector utilizes electromagnetic and hadronic interactions to capture particles. If particles are only involved in weak interactions, they leave no signatures in the detector like neutrinos or some new particles predicted by BSM theories. In this case, the only way to measure their energy is via the momentum

**Table 2.3:** b-Tagging Working Points

Cut Value	b-jet Efficiency[%]	c-jet Rejection	light-flavour-jet Rejection
0.4496	60	21	1900
-0.0436	70	8.1	440
-0.4434	77	4.5	140
-0.7787	85	2.6	28

conservation.

When two protons collide into each other in the LHC, both of them have no momentum on the transverse plane, so the transverse momentum sum of collision products is supposed to be zero. Therefore, the definition for sum of transverse momentum from invisible particles could be presented as:

$$\vec{E}_T^{miss} = \sum_{\text{visible objects}} -\vec{p}_T^i \quad (2.8)$$

where  $E_T^{miss}$  is supposed to be called missing transverse momentum, but it is called “missing transverse energy” out of historical reason. The explicit form is:

$$\vec{E}_T^{miss} = -(\sum \vec{p}_T^e + \sum \vec{p}_T^\mu + \sum \vec{p}_T^\gamma + \sum \vec{p}_T^{jet} + \sum \vec{p}_T^\tau + \sum \vec{p}_T^{soft}) \quad (2.9)$$

with  $p_T$  contributed from different objects passing loose selection. Even though tau,  $\tau$ , and photon,  $\gamma$ , are not used in the analysis, they are still reconstructed and applied in  $E_T^{miss}$  estimation. The last term is referred to the detector soft signatures which are not used in the reconstruction of any objects. They could be either from tracks (track soft term, TST) or clusters (cluster soft term, CST). The track soft term considers only the remaining tracks from the primary vertex, so it has lower dependence on the pile-up events, while it cannot deal with the contribution from neutral objects which, instead, could be recovered by CST. In ATLAS Run2 analyses, TST is preferred, because it delivers smaller uncertainty with the high pile-ups. As  $E_T^{miss}$  is only calculated on the transverse plan, it has no  $\eta$  information.

## 2.4 Simulation

For the two analyses in this thesis, the SM background estimation comes from the Monte Carlo simulation. It is performed in a couple of steps: event generation → event overlapping with “minimum bias” (MB) events → detector response simulation → digitization → physical object reconstruction → physics analysis.

Event generation is through the generators designed by theorist with the input of theoretical parameters for the interactions:  $pp(\rightarrow X) \rightarrow Y$ .  $X$  is the medium state particle with short lifetime, and it will eventually decay to  $Y$  as the final stable particles leaving signatures in the detector with the kinematic properties assigned by the event generator. The total cross section of the interaction will then be evaluated by this equation:

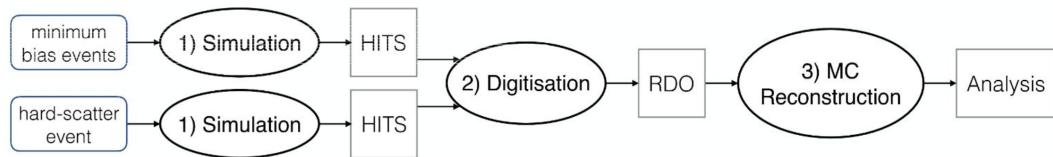
$$\sigma_{pp \rightarrow Y} = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu_1) f_b(x_2, \mu_2) \hat{\sigma}_{a,b \rightarrow X}(x_1, x_2, \mu_R) \quad (2.10)$$

with  $a, b$  as the flavours for the proton partons (quark or gluon) involved in the interaction, and  $x_1$  and  $x_2$  are the momentum fraction of the partons relative to the whole proton,  $\hat{\sigma}_{a,b \rightarrow X}$  is the cross-section calculated perturbatively of the process,  $a, b \rightarrow X$ , and  $f_a, f_b$  are the parton distribution function (p.d.f.) for the corresponding parton flavours which give the possibility of the momentum transfer from partons to the output stable particles.  $\mu_1, \mu_2$ , and  $\mu_R$  are the factorization factors which decide how the p.d.f. evolves.

For the ATLAS simulation, two processes are simulated in one event: pile-ups,  $pp \rightarrow jj$ , and hard-scattering,  $pp(\rightarrow X) \rightarrow Y$  with  $X$  and  $Y$  as the particles of interests. Generally, hard-scattering is generated with a specific generator which can give the best accuracy of simulation, and the events are then passed to *PYTHIA8* for the generation of pile-up events and the hadronizaton of the hadronic objects.

The next step is to simulate the interaction between the particles and the detector. The detector is described by *GEANT4* with the input parameters like materials and logical

volumes. The detector description is then stored in the ATLAS Geometry database which has low flexibility to change the content, while an additional database, COOL, is used to keep the information changed with time like dead channels and LAr high voltage settings. The particles are then parametrized to interact with the detector, which gives the digitalised output of raw data objects like tracks and calorimeter clusters. After the same physical object reconstruction from the detector signatures as data, the simulation samples are ready for the physical analyses in the data format called “analysis object data” (AOD). The full procedure could be seen in the diagram of Fig. 2.13



**Figure 2.13:** The full procedure of the ATLAS simulation

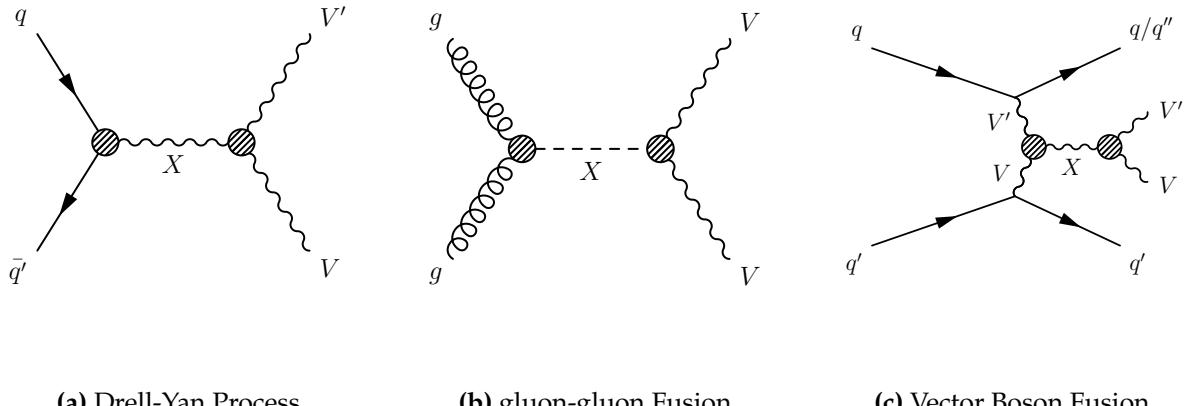


# Chapter 3

## Resonance Searching Strategy

In the study in this thesis, WW and WZ are the two medium states of interest through the production of gluon-gluon fusion, Drell-Yan process, or vector boson fusion for the BSM particles interpreted by NWA Higgs bosons, heavy vector triplets, or gravitons. The vector boson fusion is the fusion process of two vector bosons ( $W$  or  $Z$ ) emitted from two incoming quarks, and the two quarks are then scattered into two energetic jets with wide  $\eta$  separation and high invariant mass. The production processes could be seen in Fig. 3.1 as Feynman diagrams. For the final state, one  $W$  boson would decay leptonically ( $W \rightarrow l\nu$ ) into an electron or muon accompanied by a neutrino with the corresponding flavour, while the decay channel of  $\tau\bar{\tau}$  is not considered. For the other boson,  $W$  or  $Z$ , it is chosen to decays hadronically into two quarks reconstructed into two  $R = 0.4$  jets or one  $R = 1.0$  jet. ( $W/Z \rightarrow jj$  or  $W/Z \rightarrow J$ ). The benefit of choosing this final state is to have the high branch ratio from the hadronic decay and suppress the QCD contamination by the leptonic side. This study is conducted in a wide range of candidate particle mass ranging from  $300\text{GeV}$  to  $5\text{TeV}$ . If the mass of a resonance particle is high enough ( $m > 1\text{ TeV}$ ), the outcoming two quarks in the final state would be highly boosted, so they cannot not be resolved as two jets with  $R = 0.4$ , and a larger cone of  $R = 1.0$  is applied to collect their signatures into a single fat jet.

This search was performed with the  $36.1\text{fb}^{-1}$  data collected by the ATLAS detector in 2015 and 2016 with pp collisions at  $\sqrt{s} = 13\text{TeV}$ .



**Figure 3.1:** The Feynman diagrams of different production mechanisms for particle  $X$  which decays into two SM bosons.

### 3.1 Signal Models

In the SM, bosons are the force carriers and also maintain the conservation of certain physical quantities which are called symmetry. To seek for the solution of unsolved problems of the SM, many new models are predicting the existence of new bosons corresponding to unknown interactions or symmetries, and they also have the strong coupling to the SM bosons which provide the access to verify those theories. However, the existing new models are constructed with many free parameters, and each set of them needs a dedicated analysis from the experimental side, which is impossible in reality. Therefore, a simplified model with only the kinematic parameters related to resonance mass is introduced for which experiments could provide the precise measurement to on-shell bosons.

This strategy could scan through many models, so it is defined as a general search. However, to give a better separation between signal and background, three benchmarks are applied in this analysis for sensitivity optimization which corresponds to bosons with different spins:  $spin = 0$ , narrow width approximation Higgs bosons (NWA);  $spin = 1$  heavy vector triplets (HVT);  $spin = 2$ , Randall–Sundrum model gravitons (RSG).

### Narrow Width Approximation Higgs Boson

With the problem in Higgs boson naturalness, some extended models predict the existence of high mass Higgs bosons. However, as only the kinematic properties are concerned, the interpretation model chosen in this analysis is the SM Higgs boson but with higher mass. To have further simplification, the decay of the Higgs boson is forced to be always at the mass pole with the narrow width approximation. This means the transferred,  $q$ , from the proton partons is exactly the mass of the resonant particle under the assumption which gives the narrow resonance width,  $\Gamma/m_H \ll 1$ , and the interference to the SM Higgs boson is taken negligible. Therefore, the Relativistic Breit–Wigner distribution could be written as:

$$f(q) = \frac{k\pi}{M\Gamma} \delta(q^2 - m_H^2) \quad (3.1)$$

where  $k$  represents:

$$k = \frac{2\sqrt{2}m_H\Gamma\gamma}{\pi\sqrt{m_H^2 + \gamma}} \quad (3.2)$$

and  $\gamma = \sqrt{m_H^2(m_H^2 + \Gamma^2)}$ . This is then used to evaluate the cross-section of the Higgs boson production.

### Heavy Vector Triplet

Heavy vector bosons are predicted by many new BSM theories with the coupling to quarks, leptons, SM vector bosons and Higgs bosons, which constructs a wide phase space to explore. To examine the suitable theories, this study attempts to investigate all the couplings with the set-up of one neutral heavy boson,  $Z'$ , and two degenerate charged bosons,  $W'^{\pm}$ , with the given coupling constant,  $g_V$ . For optimization, two models are taken as the benchmarks. Model A is with an additional symmetry breaking to SM,  $SU_1(2) \times SU_2(2) \times U(1) \rightarrow SU_L(2) \times U(1)$  giving a weak coupling:  $g_V \sim \mathcal{O}(1)$ . For the scenario of a strong SM boson coupling, the Minimal Composite

Higgs Model is taken as model B with the symmetry breaking,  $SO(5) \rightarrow SO(4)$  for  $4\pi \geq g_V \geq 1$ . However, because the decay width is proportional to the coupling constant, and the focus of this search is for the narrow resonance, only  $6 \geq g_V \geq 1$  is considered with  $\Gamma_{V'}/m_{V'}$  below 10%.

To simplify the models, the coupling strength to all fermions are equal with the scale of  $g^2 c_F/g_V$  where  $g$  is the  $SU_L(2)$  gauge coupling, and  $c_F$  is the dimensionless coefficient between bosons and fermions defined as a free parameters of order one in the phase space of interest. As the fermionic coupling scale is proportional to  $1/g_V$ , model A turns to be more sensitive to the fermionic production with Drell-Yan process, while model B is well-suppressed. In the contrary, the coupling to bosons is governed by  $c_H g_V$  with  $c_H$  as the universal coupling among bosons. Therefore, model B has better sensitivity for higher branch ratio of the decay channel of diboson in this analysis than model A. For the interpretation, the two parameters,  $g^2 c_F/g_V$  as well as  $c_H g_V$ , construct a two-dimension phase space, and it could be explored with production rate and decay branch ratio.

As the coupling to all bosons are the same ( $c_H g_V$ ), the neutral and charged heavy boson ( $Z'$  and  $W'^{\pm}$ ) have the same decay branch ratio to all SM bosons:

$$BR(Z' \rightarrow ZH) = BR(Z' \rightarrow W^\pm W^\pm) = BR(W'^\pm \rightarrow W^\pm Z) = BR(W'^\pm \rightarrow W^\pm H) \quad (3.3)$$

However, with the small mixing angle (between SM and BSM bosons), the coupling in the transverse component is well suppressed, and the dominant contribution is from longitudinal component. For the same reason, the coupling to neutral dibosons are also so weak that those channels are ignored in this analysis, which is also applicable to  $W\gamma$ . For the coupling to  $HH$ , they are forbidden due to the concern of momentum and angular momentum conservation.

### Randall-Sundrum Graviton

To solve the hierarchy problem, extra dimensions were proposed as one of the so-

lutions. It leads to the result that the observed Planck scale,  $M_{pl} = 2 \times 10^8 GeV$ , is already reduced due to the existence of extra dimensions from the orginal scale,  $M$ . The relation between  $M_{pl}$  and  $M$  is:

$$M_{pl}^2 = M^{n+2} V_n \quad (3.4)$$

where  $n$  is the number of dimensions which are not yet observed, and  $V$  is the volume constructed from the extra dimensions regardless of the four-dimensional spacetime. Therefore, the visible spacetime is just a manifold under  $(4 + n)$  dimensions.

Under Randall-Sundrum model, only one more dimension is needed, which hypothesizes that the fifth dimension is constrained with boundary condition of the  $\phi$  periodicity ranged between  $-\pi$  to  $\pi$  called the “warped bulk” which bridges two four-dimensional manifolds at  $\phi = \pi$  and  $\phi = 0$  ( $\phi$  is taken as the fifth coordinate). The “Hilbert-Einstein”action under the set-up could be presented as:

$$S = S_{gravity} + S_{obs} + S_{hid} \quad (3.5)$$

$$S_{gravity} = \int d^4x \int_{-\phi}^{\pi} d\phi \sqrt{G} \left[ -\Lambda + 2M^3 R \right] \quad (3.6)$$

$$S_{vis(hid)} = \int d^4x \sqrt{-g_{vis(hid)}} \left[ \mathcal{L}_{vis(hid)} - V_{vis(hid)} \right] \quad (3.7)$$

with  $\Lambda$  as the cosmological constant,  $R$  as the scalar spacetime curvature, and  $g$ 's are the determinants of metric tensor matrix,  $g_{\mu\nu}$ ,  $V_{vis}$ , and  $V_{hid}$  are the constant gravitational potentials taken out from the Lagrangian vacuum energy for the visible and hidden spacetimes. After inserting the terms into the Einstein Feild Equation, it leads to the solution for the spacetime description:

$$ds^2 = e^{-2\sigma(\phi)} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2 \quad (3.8)$$

with

$$\sigma(\phi) = kr_c|\phi| \quad k = \sqrt{\frac{-\Lambda}{24M^3}} \quad (3.9)$$

where  $\eta$  is the Minkowski metric, and  $r_c$  is the constant independent of  $\phi$  taken as the “compactification radius” of the extra dimension on the orbifolding. As a result, the extra dimension only has the dimensional interval,  $\pi r_c$ , at  $\phi = \pi$  in the visible spacetime. Taking the space description into Eq. 3.4, the relation between  $r_c$  and  $M_{pl}$  could be derived as:

$$M_{pl}^2 = \frac{M^3}{k} \left[ 1 - e^{-2kr_c\pi} \right] \quad (3.10)$$

This expression indicates that  $M_{pl}$  depends on  $kr_c$ , and the weak gravity could be explained with a proper choice of  $r_c$ . Under the solution, the existence of graviton (the gravitational field) is then taken as the tensor fluctuation on Minkoski metric:  $\eta_{\mu\nu} \rightarrow \eta_{\mu\nu} + \bar{h}_{\mu\nu}(x)$ . To estimate its mass, the new spacetime geometry is inserted into the Higgs sector in the SM Lagrangian, and it gives the result:  $m = e^{-kr_c\pi} m_0$  with  $m_0$  as the original mass scale in the visible manifold (IR brane), and  $m$  as the one in the five-dimensional spacetime. (This relation could also be applied to SM particles.) If  $e^{kr_c\pi}$  is of the order  $10^{15}$ , the mass scale would be in the scale of  $TeV$  under the mechanism which offers the signature verifiable to the LHC energy scale with the couplings to SM particles derived from the same way.

## 3.2 Simulation Samples and Derivation

Each SM background process and mass points of signals are simulated by the procedure mentioned in 2.4. To make a proper comparison between the simulation and data, the event numbers are normalised to the theoretical cross section and total data luminosity. However, the modelling of interactions between the ATLAS detector and particles is not perfect, and it leads to the discrepancy in efficiency measurements including the particle reconstruction, lepton isolation, trigger, and jet b-tagging effi-

ciency. To recover this disagreement, scaling factors are estimated from the comparison between data and MC and applied on the event weight in the MC samples.

In addition, the other disagreement comes from the inconsistency in distribution of interaction number per bunching crossing,  $\mu$ . To eliminate the effect, another scale factor is applied through the process called “pile-up reweighting” (PRW) to make the simulated  $\mu$  distribution agree with data.

After considering all the factors for the data-MC comparison, the final simulation event yield could be reweighted to data by:

$$N_{yield} = \mathcal{L} \times XS \times \epsilon_{rec} \times \epsilon_{iso} \times \epsilon_{trigger} \times \epsilon_{b-tagging} \times \epsilon_{prw} / N_{mc} \quad (3.11)$$

where  $N_{mc}$  is the total event weight from simulation, and  $\epsilon$ 's stand for the scaling factors of different contributions.

### Background Simulation

Some of the SM processes have the same final state to the new physics of our interest: one lepton, one neutrino and multiple jets, and they are called “irreducible” background which could not be well-suppressed by selection cuts. This type of backgrounds are estimated from the Monte Carlo simulation contributed from W+jets ( $W \rightarrow l\nu$ ),  $t\bar{t}$  ( $t \rightarrow bW \rightarrow bjj$  and  $t \rightarrow bW \rightarrow bl\nu$ ), diboson ( $WW/WZ \rightarrow l\nu jj$ ), Z+jets ( $Z \rightarrow ll$ ), and single top interactions.

The events of W/Z+jets are simulated by SHERPA v2.2.1, with the PDF configuration of NNPDF30NNLO as the baseline generator, and the simulation uncertainty is taken by the comparison to other generators detailed in next chapter. With the complicated process of hadronisation including the broad range of jet  $p_T$  and involved quark flavours, the simulation is done respectively with multiple slices of  $\max(h_T, p_T(W/Z))$  ( $h_T$  is the scalar sum of  $p_T$  from all jets) and different number of bottom and charm quarks. The involved matrix element for the simulation are up to 2

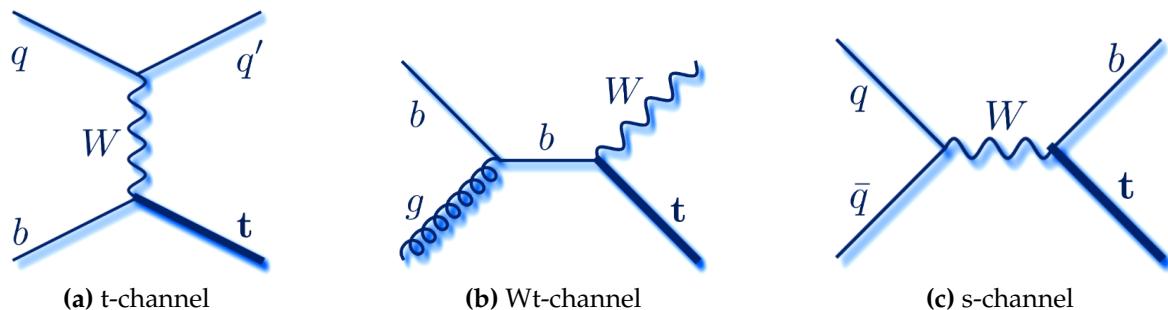
partons at NLO and 4 partons at LO which is followed by merging into the Sherpa parton shower. The resulting cross section for normalisation is estimated to NNLO of QCD.

$t\bar{t}$  events are generated through POWHEG-BOX v2 with the matrix element calculation provided by CT10 PDF with top quark mass set at  $172.5\text{GeV}$ , and the HDAMP parameters for high  $p_T$  radiation is set at  $1.5m_t$ . Different from SHERPA as a self-contained generator to do parton shower itself, the simulation from POWHEG-BOX is then interfaced through MADSPIN and PYTHIA8.186 tuned by Perugia 2012 (P2012) and CTEQ6L1 PDF sets for spin correlation preservation of top quark decays and the following parton shower, fragmentation and underlying events. The renormalisation and factorisation scale of the whole process are determined by  $\sqrt{m_t^2 + p_T^2(t)}$ . The  $t\bar{t}$  cross section used for normalisation is calculated using TOP++ 2.0 with the precision up to NNLO in QCD. To take in the contribution from soft gluon terms, a re-summation with next-to-next-to-leading logarithmic (NNLL) is applied to make further correction.

Single top events are generated through three processes: s-, t- and Wt-channel productions (Feynman diagrams are presented in Fig. 3.2). For the simulation of Wt and s-channels, the same recipe from  $t\bar{t}$  generation is adopted, while the t-channel one is through POWHEG-BOX v1 with fixed four-flavor CT10f4 PDF set but also followed by the same procedure for decay and parton showering from  $t\bar{t}$  generation. The renormalisation and factorisation scales are set respectively for the three channels with:

- s-channel & Wt-channel:  $m_t$
- t-channel  $4 \times \sqrt{m_q^2 + p_T^2(q)}$  ( $q$  is the quark associated with the single top quark production)

, and the cross section for each production is calculated separately with the description in



**Figure 3.2:** The Feynman diagrams of three channels for single top production.

The generation of WW/WZ events are also through SHERPA v2.2.1 for the event production and the hadronisation.

## Signal Simulation

HVT samples are generated via MADGRAPH5 interfaced to PYTHIA8 with the resonance mass points ranged from  $300\text{ GeV}$  to  $5\text{ TeV}$  of  $100\text{ GeV}$  spacing. For simplicity,  $g_V = 1$  and  $g_V = 3$  are set for model A and model B respectively.

RS graviton events are also simulated through MADGRAPH5 and PYTHIA8, and only the ggF production is considered for this signal. Within the simulation,  $r_c = 1$  is set as the default for the simulation, but it is also reweighted in the resonance mass distribution at parton level for  $r_c = 0.5$ . This is for the comparison with the result from the CMS collaboration. The decay width of this configuration is expected to be  $\approx 6\%$ .

The decay width and cross section of HVT and RS graviton are summarised in tab. 3.1: For the NWA Higgs boson, its interference to the SM Higgs boson ( $125\text{GeV}$ ) is

assumed to be negligible as discussed in 3.1. Its narrow decay width is set as a constant at  $4.07\text{MeV}$  for all mass points which is beyond the experimental resolution with the production of ggF and VBF, which are simulated separately. The simulation is done by POWHEG-BOX v2 showered with PYTHIA8 under CTEQ6L1 PDF set.

## Derivation

**Table 3.1:** The decay width and cross section of HVT and RSG at  $800GeV$ ,  $1.6TeV$ , and  $2.4TeV$  mass points

$m$ [TeV]	HVT $W'$ and $Z'$			RS $G^*$	
	$\Gamma$ [GeV]	$\sigma \times BR(Z' \rightarrow WW)$ [fb]	$\sigma \times BR(W' \rightarrow WZ)$ [fb]	$\Gamma$ [GeV]	$\sigma \times BR(G^* \rightarrow WW)$ [fb]
0.8	32	354	682	46	301
1.6	51	38.5	79.3	96	4.4
2.4	74	4.87	10.6	148	0.28

To boost the computing power, the analyses are not operated on AODs directly, and, instead, they will be going through the “derivation” procedure composed of “trimming” and “slimming” to drop down variables and events of no interest first, which outputs the data format called derived AOD (DAOD). For the broad variety of analysis types, a couple of derivation schemes are applied, and the analyses with similar final states share the same derivation scheme.

With the final state of this analysis, “HIGG5D2” is chosen with the derivation scheme as the following:

- trigger: passing at least one electron, muon, or  $E_{missing}^T$  trigger
- lepton: one electron or muon with  $p_T > 15GeV$
- jet: two small R jets with  $p_T > 20GeV$ , one small R jet with  $p_T > 100GeV$ , or one large R jet with  $p_T > 150GeV$

### 3.3 Physical Object Definition

Because the LHC is using protons as the beam source, it leads to the enormous production of hadronic jets. Within the environment, most of other objects suffered great contamination from jet misidentified as other objects. Therefore, the object definition for this analysis is to keep the signal efficiency and significant suppression of misidenti-

fication of the intended objects at the same time.

## Electron

The electrons in this analysis are defined as two types, loose and signal, and each event only has exactly one signal lepton without additional loose one. Signal electrons are required to have  $p_T$  above 27 GeV to reach the trigger efficiency turn-on plateau, and  $|\eta| < 2.47$  is applied on both electron types within the acceptance of inner detector with the crate region vetoed ( $1.37 < |\eta| < 1.52$ ). The impact parameter requirement is set to only consider the electrons from the primary vertex. The full selection criteria is shown in Tab. 3.2.

In addition to the fundamental quality requirement, the overlap removal is applied afterwards to prevent the objects reconstructed from the same detector signature. When an electron shares inner detector tracks with any muon candidate, the electron is discarded. The existence of a nearby jet defined by  $0.2 < \Delta R(e, j) < \min(0.4, 0.04 + 10/p_T(e)[GeV])$  also makes the electron removed. The final requirement on electron is that it shall be consistent with the trigger level electron which fired the required electron trigger to suppress the QCD background.

**Table 3.2:** Selection for electron candidates used in the analysis. Veto and signal electrons are defined.

	Electrons	
	Loose	Signal
$p_T$	$> 7\text{GeV}$	$> 27\text{GeV}$
$ \eta $	$< 2.47$	$\notin [1.37, 1.52]$
Identification	LooseLH	TightLH
Isolation	LooseTrackOnly	FixedCutTight
$ d_0/\sigma(d_0)^{BL} $	$< 5$	
$ z_0 \sin \theta $	$< 0.5 \text{ mm}$	

## Muon

Similar to electrons, loose and signal muons are defined with  $p_T$  and  $|\eta|$  cuts in the consideration of trigger turn-on curve plateau and inner detector coverage. The requirement on muon impact parameters is tightened for better rejection to the cosmic muons. The selection criteria is shown below in Tab. 3.3

As muons have the lowest misidentification rate, they are kept in most cases for overlap removal. The only exception is the muons decayed from heavy flavour quark, which are called non-prompt muons. To remove this type of contamination, the muons are discarded under the scenarios:

- $\Delta R(\mu, j) < 0.2$
- $\Delta R(\mu, j) < \min(0.4, 0.04 + 10/p_T(\mu)[GeV])$

with the jets fulfilling either of the conditions: a)  $p_T^\mu / p_T^j < 0.5$  and number of jet-associated tracks greater than 2, b)  $p_T^\mu / \sum_1^n p_T^{trk} < 0.7$  for all the jet-associated tracks and  $n > 2$ .

The last selection in muon is that it shall be spatially consistent to the trigger muon if muon trigger is fired in the event.

**Table 3.3:** Selection for muon candidates used in the analysis. Veto and signal electrons are defined.

	Muons	
	Loose	Signal
$p_T$ threshold	7 GeV	27 GeV
$ \eta $	$< 2.7$	$< 2.5$
Identification	Loose	Medium
Isolation	LooseTrackOnly	FixedCutTightTrackOnly
$ d_0/\sigma(d_0) w.r.t.BL$		$< 3$
$ z_0 \sin \theta $		$< 0.5$ mm

### Small R Jets [R=0.4]

In the intended final states, the jets (denoted as  $j$ ) come from the decay of W bosons

$(W \rightarrow jj)$  or the remnant quarks from the vector boson fusion ( $jj \rightarrow WWjj$  or  $jj \rightarrow WZjj$ ). Because of the kinematic properties, the two types of jets are selected respectively. The full selection criteria are in Tab. 3.4.

The pair of VBF jets are supposed to be a high mass dijet system with wide separation, so they have tighter  $p_T$  selection of  $p_T > 30\text{GeV}$  but looser  $|\eta|$  cut,  $|\eta| < 4.5$ . For signal jets, they are only required to have  $p_T > 20\text{GeV}$ , and only the ones within the acceptance of inner detector ( $|\eta| < 2.5$ ) are taken as jet candidates for event selection. The jet quality requirement is to remove the “fake jets” from calorimeter noise pulse, cosmic ray, or non-collision background (like beam-halo), which is called “jet cleaning”.

**Table 3.4:** Selection for small-R jets

	Small-R Jets	
	Signal Jets	VBF Jets
Algorithm	anti- $k_t$ , $R = 0.4$	
$p_T$	$> 20\text{GeV}$	$> 30\text{GeV}$
$ \eta $	$< 2.5$	$< 4.5$
Quality	not “bad” jet	
JVT	$< 0.59$ ( $ \eta  < 2.4$ && $p_T < 60\text{ GeV}$ )	
b-Tagging	MV2c10, 85% efficiency	

### Large R Jets [R=1.0]

When the W or Z boson is highly boosted decayed from a heavy particle, the outgoing quarks would be close to each other. In this case the small R jets would not have enough resolution power to reconstruct them individually, so the large R jets (or called “fat jets” and denoted as  $J$ ) are reconstructed to collect the energy deposits from the close-by quarks. The full selection on the fat jets could be seen in Tab. 3.5. With this topology, the jet mass and  $p_T$  would need the further correction. This is performed with the track-assisted mass,  $m^{TA}$ , as the calorimeter cannot provide enough spatial resolution.  $m^{TA}$  is estimated from the tracks left by charged jet partons inside the fat

jets defined as:

$$m^{TA} = m^{trk} \times \frac{p_T^J}{\sum p_T^{trk}} \quad (3.12)$$

Here,  $m^{trk}$  is the reconstructed mass of the tracks taken as massless particles, and  $p_T^{trk}$  is the vector sum from  $p_T$  of tracks. The ratio of  $p_T$  between tracks and the jet is to take in the neutral-to-charge fluctuations. It could then be combined with the calorimeter mass,  $m^{calo}$ , into the combined mass,  $m^{comb}$ , by this definition:

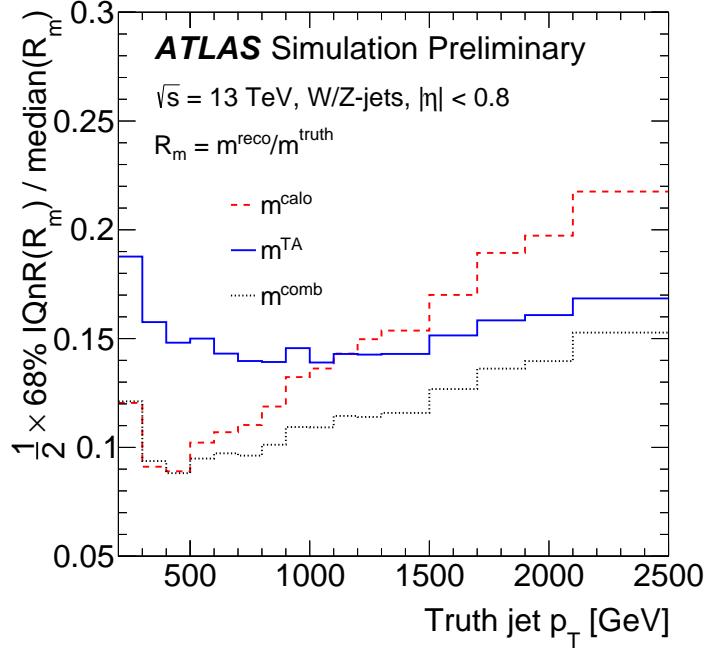
$$m^{comb} = \frac{\sigma_{calo}^{-2} m^{calo} + \sigma_{TA}^{-2} m^{TA}}{\sigma_{calo}^{-2} + \sigma_{TA}^{-2}} \quad (3.13)$$

with  $\sigma_{calo}^{-2}$  and  $\sigma_{TA}^{-2}$  as pre-estimated mass resolutions for the calorimeter and track-assisted mass which are assumed to be uncorrelated. From Fig. 3.3, it could be seen that the calorimeter mass has better performance in the low  $p_T(W)$  regime benefited from the great energy resolution, but it is degraded as  $p_T(W)$  increases, while the track-assisted mass performed in an opposite way. The combined mass takes the merits of these two mass definitions and provide the best mass resolution ( $\tilde{10}\% (15\%)$  at jet  $p_T = 1\text{TeV}(2.5\text{TeV})$ ). It is taken as the nominal fat jet mass in this analysis with the selection of  $m^{comb} > 50\text{GeV}$ . The jet  $p_T$  is the corrected by  $p_T^{comb} = p_T^{calo} \times m^{comb} / m^{calo}$

**Table 3.5:** Selection for large-R jets

	Signal Large-R Jets
Algorithm	anti- $k_t$ , $R = 1.0$
$p_T$	$>200\text{ GeV}$
$ \eta $	$< 2.0$
Mass threshold	50 GeV
W/Z Tagger	$D_2^{\beta=1} \& m^{comb}$

However, the combined mass is still not proficient to select the W/Z decayed fat jets precisely, so the substructure of jets is needed to improve the boson tagging. This extra information is extracted with the subjets of  $R = 0.2$  from a  $k_T$  algorithm performed



**Figure 3.3:** The jet mass resolution as a function of jet  $p_T$  for jets produced from boosted  $W$  boson [?]. Three different jet mass reconstruction algorithms are displayed: the calo-jet mass ( $m^{\text{calo}}$ ), the track-assisted mass ( $m^{\text{TA}}$ ), and the combined TA+calo mass ( $m^{\text{comb}}$ ).

on the topoclusters. Those tiny jets are then taken as the new entities to be “ghost-associated” with the fat jets through the  $\text{anti}-k_T$  algorithm with the threshold on  $p_T$ ,  $p_T^{R=0.2}/p_T^{R=1.0} > 0.05$ . The jet substructure information could then be give by the discriminant,  $D_2^{\beta=1}$ , for  $W/Z$  boson recognition which is defined as:

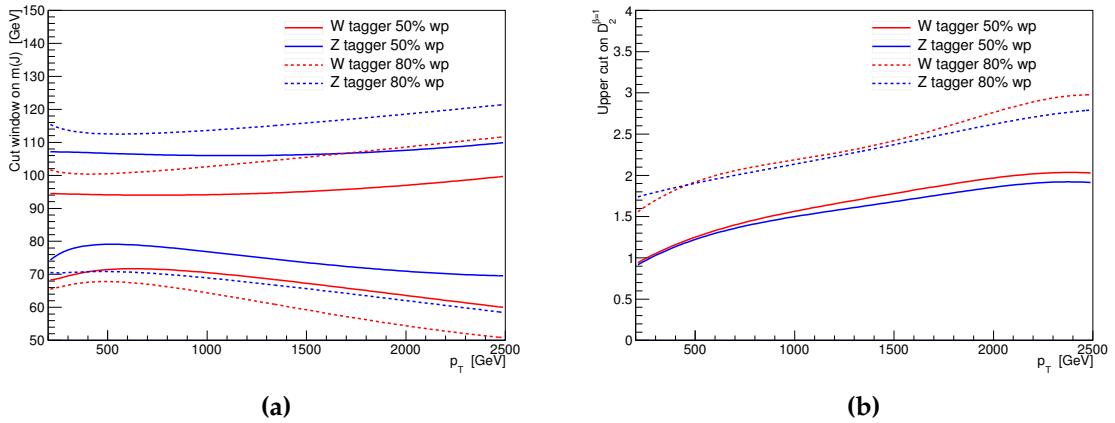
$$D_2^{\beta=1} = \frac{e_3^\beta}{e_2^\beta} \quad (3.14)$$

with  $e_2^\beta$  and  $e_3^\beta$  as:

$$e_2^\beta = \frac{1}{(p_T^{\text{jet}})^2} \sum_{i < j \in J} p_T^i p_T^j (R_{ij})^\beta \quad (3.15)$$

$$e_3^\beta = \frac{1}{(p_T^{\text{jet}})^3} \sum_{i < j < k \in J} p_T^i p_T^j p_T^k (R_{ij} R_{jk} R_{ik})^\beta \quad (3.16)$$

where  $i$ ,  $j$ , and  $k$  are the index for the subjets. The boson tagging is then done by a 2D cut on both  $D_2^{\beta=1}$  and  $m^{comb}$  as a function of  $p_T$  shown in Fig. 3.4 with two working points, 50% and 80%, for the tagging efficiency.



**Figure 3.4:** The thresholds of the mass window cut (a) and the upper cut on  $D_2^{\beta=1}$  (b) as a function of  $p_T$  used in this analysis. The cuts for  $W$ -( $Z$ )-boson tagging is shown by red (blue) lines.

## Missing Transverse Energy

Although  $E_T^{missing}$  is supposed to be reconstructed as shown in Sec. 2.3, tau and photons are treated as jets for the intended final state in this analysis.

The cut on  $E_T^{miss}$  will be discussed in the next section.

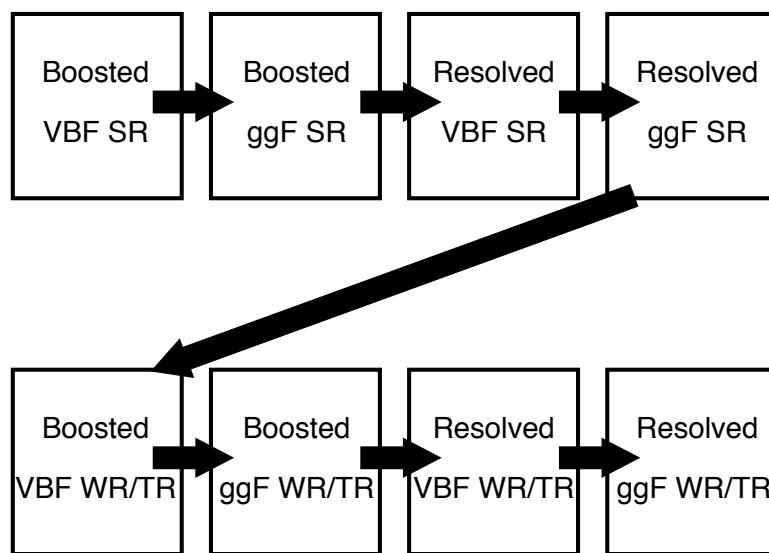
## 3.4 Event Selection

The event selection is conducted to maximize the sensitivity by removing background events but also keeping the signal at the same time. To achieve this purpose, the significance is defined as:

$$\sigma_{sig} = \sum_i^{N_{bin}} \left( \frac{s_i}{s_i + b_i + (\Delta_b i)^2} \right)^2 \quad (3.17)$$

where  $s_i$  and  $b_i$  are the signal and background event numbers. The cuts on variables would be varied on signal and background samples simultaneously, and the final criteria is given by the combination of cuts giving the best significance. The signal sample applied in the optimization could be either via the medium state of WW or WZ, as they have similar kinematic properties. However, they are still divided into two subchannels with the definitions for dedicated mass windows.

Then, the events in this analysis are further categorized by jet topologies and VBF jet selection. In general, the VBF categories could gain better sensitivity than the ggF/DY production ones, and the events with fat jets which are called “boosted” events are also more sensitive than the resolved signal regions. Fig. 3.5 shows how the events are categorized, and the categories with better sensitivity to the signal are given high priorities (boosted > resolved, VBF > ggF/qqF). For the events which fail part of the jet selection, they would go into the control regions to constrain the background contribution from  $W+jets$  and  $t\bar{t}$  which are the two dominant backgrounds in this analysis. The two control regions are used in the simultaneous fitting to derive the scale factor of background estimation, and the details will be discussed in the next chapter



**Figure 3.5:** Illustration of how to combine the SR/CRs in this analysis.

### 3.4.1 Trigger

The first applied criterion on event selection is trigger. The recorded data is a broad collection of different physical signatures, and the final state in the analysis only accounts for a small fraction of them. Therefore, certain triggers which could be fired by the signatures from our final state is required, which are the electron, muon, and  $E_T^{miss}$  triggers. Because of the increasing luminosity provided by LHC, the trigger threshold was enhanced in 2016 to reduce the trigger rate. For the MC samples, run number is randomly generated, and the events shall only pass the triggers available in the random run number. The full trigger set used in this analysis is shown in Tab. 3.6 Three electron triggers are used in electron channel including the unprescaled

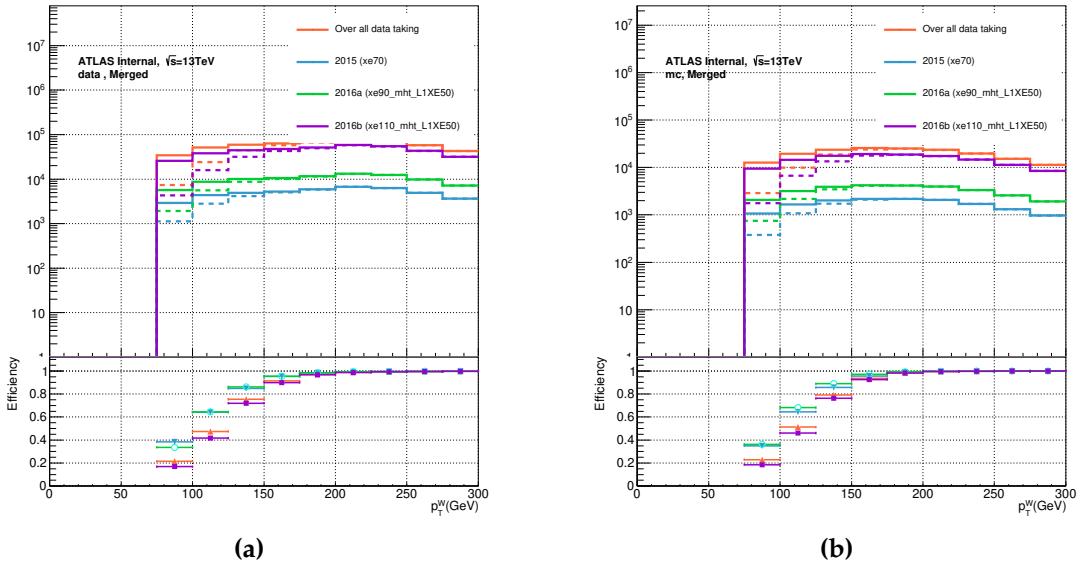
**Table 3.6:** The list of triggers used in the analysis.

Data-taking period	Electron channel	Muon channel	
		$p_T, (\mu\nu) < 150\text{ GeV}$	$P_T, (\mu\nu) > 150\text{ GeV}$
2015	HLT_e24_lhmedium_L1EM20 HLT_e60_lhmedium HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50	HLT_xe70
2016a (run < 302919) ( $L < 1.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ )	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50	HLT_xe90_mht_L1XE50
2016b (run $\geq 302919$ ) ( $L < 1.7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ )	same as above	same as above	HLT_xe110_mht_L1XE50
Total int. lumi. [ $\text{fb}^{-1}$ ]	36.1	35.6	35.9

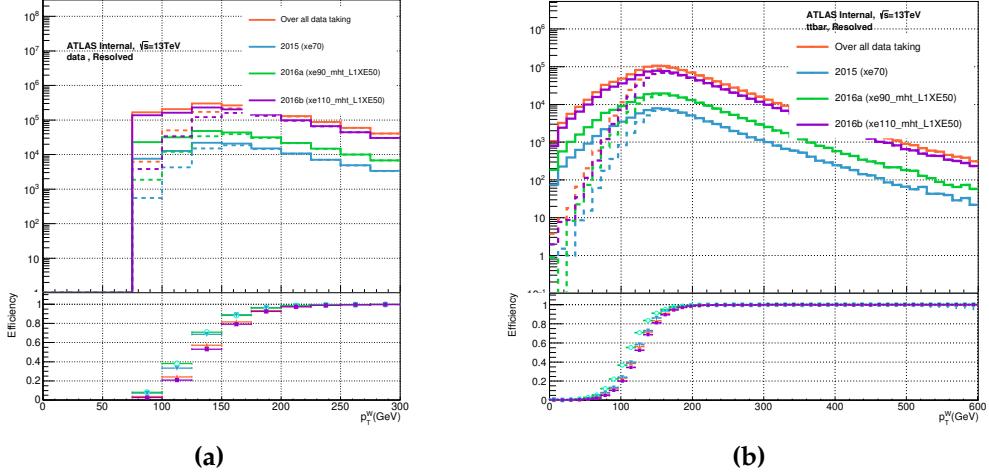
lowest threshold one to maximize the signal efficiency. The other two triggers are used to select high  $p_T$  electrons with looser isolation requirement. The combined performance of the triggers is around 90% efficiency at the turn-on plateau as a function of  $p_T$ .

In muon channel, both  $E_T^{miss}$  and muon triggers are used. For the scenario of  $p_T(\mu\nu) < 150\text{GeV}$ , the unprescaled lowest threshold muon trigger is used accompanied by the higher threshold one without any isolation requirement. Otherwise,  $E_{miss}^T$  trigger is chosen for  $p_T(\mu\nu) > 150\text{GeV}$  events because muon trigger can only reach 70% efficiency on the plateau.

However, the  $E_T^{miss}$  cut in this analysis is below the plateau, so there might be the inconsistency between data and simulation in terms of the efficiency. Therefore, “tag and probe” method is applied to study the trigger efficiency as a function of  $p_T(\mu\nu)$  (because muons are invisible in trigger level, trigger level  $E_T^{miss}$  is actually  $p_T(\mu\nu)$ ). This study is performed on boosted and resolved channels respectively. The tagged events are required to fulfil the following conditions for the resolved (boosted) channel: a) one muon with  $p_T > 27\text{GeV}$  b)  $E_T^{miss} > 60(100)\text{GeV}$  c) at least 2 signal jets (1 fat jet) selected d) the unprescaled lowest threshold muon trigger is fired. The events are then probed by whether  $E_T^{miss}$  trigger is fired to give the efficiency. The result for data and simulated  $t\bar{t}$  events are shown in Fig. 3.6 for boosted channel in Fig. 3.7 for the resolved channel. The efficiency reaches the plateau at  $200\text{GeV}$ , but  $E_T^{miss}$  trigger is applied for  $p_T(\mu\nu) > 150\text{GeV}$ , so the scaling factor is taken into simulation as the extra event weight to make them consistent.



**Figure 3.6:** The upper plot is  $p_T(\mu\nu)$  distribution of tagged (real) and probed events in boosted channel for data (a) and  $t\bar{t}$  events (b). The lower plots is the efficiency as a function of  $p_T(\mu\nu)$



**Figure 3.7:** The upper plot is  $p_T(\mu\nu)$  distribution of tagged (real) and probed events in resolved channel for data (a) and  $t\bar{t}$  events (b). The lower plots is the efficiency as a function of  $p_T(\mu\nu)$

### 3.4.2 Event Cleaning and Preselection

After the trigger, the event “quality is verified by a series of flags in data determining the suitability of an event for physical analyses. The following is the list:

- **Good Run:** if the detector operates in a proper status without intolerable defects, they go into the good run list (GRL). Only the events contained in the GRL are considered in this analysis.
- **Primary Vertex:** because all the physical objects are required to origin from the primary vertex, its existence is essential. Events without a proper primary vertex (defined in Sec. 2.3) are discarded.
- **Tile Error Veto:** part of the channels in tile detector are broken. If they accept any physical objects, this flag would be marked, and the events are vetoed.
- **LAr Error Veto:** part of the channels in LAr detector are broken. If they accept any physical objects, this flag would be marked, and the events are vetoed.
- **SCT Error Veto:** part of the channels in SCT detector are broken. If they accept any physical objects, this flag would be marked, and the events are vetoed.

- **Core Error Veto:** during data-taking periods, the atlas central DAQ system might suffer from some glitches which broke the data recording, and the flag is marked for events. They are also vetoed in this analysis.

### 3.4.3 Parameter of Interest, $m_{WV}$

This analysis is searching for the mass resonance of exotic particles, so it is the discriminant to seek for the signal. (i.e.  $m_W V$  distribution is the input for statistic interpretation.) However, the longitudinal  $p_T$  of neutrinos in the final state could not be measured, so the mass resolution is poor to spot the signal spike. Therefore,  $p_z(\nu)$  is solved with the assumption that  $W \rightarrow l\nu$  is the  $E_T^{miss}$  contribution in all the events although it is not held true every event, so the equation of energy conservation of W boson decay could be written down as:

$$m_W^2 = m_l^2 + 2E_l \sqrt{p_{T,\nu}^2 + p_{z,\nu}^2} - 2\vec{p}_{T,l} \cdot \vec{p}_{T,\nu} - 2p_{z,l}p_{z,\nu} \quad (3.18)$$

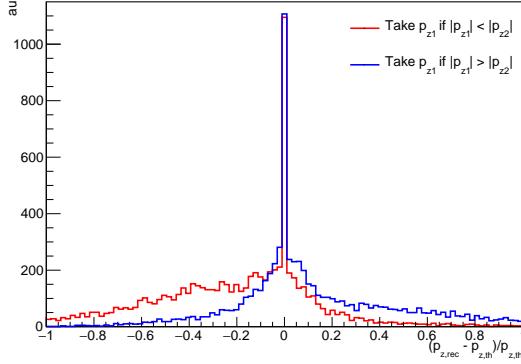
In SM, W bosons have the mass of  $80\text{GeV}$ , so  $m_l$  for electrons and muons is negligible. This leads to the quadratic equation of  $p_z^\nu$ :

$$4p_{T,l}^2 p_{z,\nu}^2 - 4(m_W^2 + 2\vec{p}_{T,l} \cdot \vec{p}_{T,\nu}) p_{z,l}p_{z,\nu} - (m_W^2 + 2\vec{p}_{T,l} \cdot \vec{p}_{T,\nu})^2 + 4p_l^2 p_{T,\nu}^2 = 0 \quad (3.19)$$

If the solutions are complex numbers, only the real terms are taken into this analysis. To determine which solution to use, the resolution is compared with the absolute value of solutions (bigger one and smaller one) defined as:

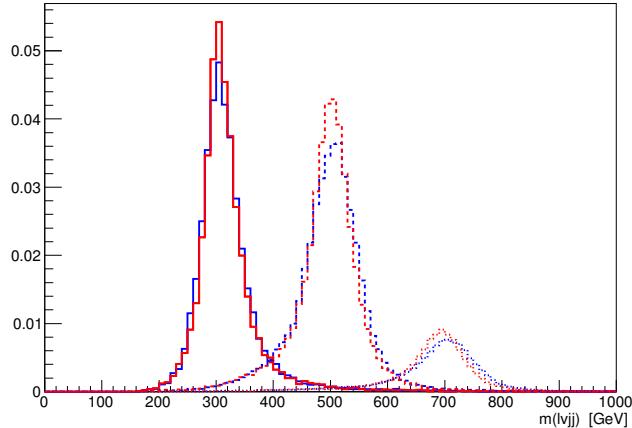
$$\sigma = \frac{p_z^{truth} - p_z^\nu}{p_z^{truth}} \quad (3.20)$$

with  $p_z^{truth}$  as the neutrino longitudinal momentum at generator level (MC truth). The result could be seen in Fig. 3.8, and it indicates the bigger one has slightly better performance in terms of the mass resolution, so it is kept. In addition to the correction on the leptonically decayed W boson, the other further improvement on  $m_{WV}$  reconstruction is also made by  $p_T(j,j)$  of the two resolved signal jets by the



**Figure 3.8:** The  $p_z^V$  resolution with absolute values of the solutions, bigger and smaller one.

correction of  $p_T^{corr} = p_T(j,j) \times \frac{m_V}{m_{jj}}$  (For WW,  $m_V$  is taken as W boson mass, while it is Z boson mass for WZ medium state). The improvement of the “mass-constraint” correction could be seen in Fig. 3.9 for  $\approx 20\%$  better  $m_{WW}$  resolution.



**Figure 3.9:**  $m_{WW}$  distributions for  $gg \rightarrow H \rightarrow WW$  signals at  $m = 300\text{GeV}$  (solid),  $500\text{GeV}$  (dashed) and  $700\text{GeV}$  (dot), with (red) and without (blue) W-mass constraint to  $W \rightarrow jj$  system.

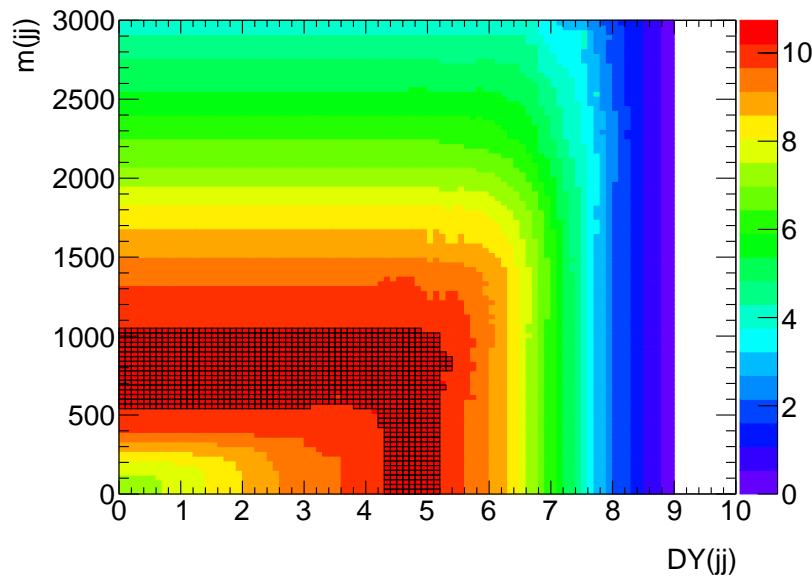
### 3.4.4 VBF Event Selection

As VBF signal regions have better sensitivity than ggF/DY ones, the selection criteria play an important role in this analysis. The optimization on the selection is conducted in three steps. First, all VBF events are required to have at least 4(2) jets in the resolved

(boosted) channel. Second, the two jets in the dijet system are supposed to be the pair with the highest mass, opposite  $\eta$  signs, and not b-tagged. This pair was chosen prior to the  $W/Z \rightarrow jj$  signal jet selection and removed from signal jet candidates. Then, the optimization is performed on a 2-dimensional phase space constructed by  $\Delta\eta(j,j)$  and  $m(j,j)$  which are two most evident signatures of this production process. The performance of the cuts on the two variables is determined by signal significance in Eq. 3.17. Fig. 3.10 shows the result of the optimization performed on the signal sample with  $700\text{GeV}$  HVT, and the best significance could be achieved by:

- $m_{jj}^{\text{VBF}} > 770\text{GeV}$
- $\Delta\eta(jj) > 4.7$

The other reason to choose this set of cuts is to make it consistent with  $WZ/ZZ \rightarrow lljj/\nu\nu jj$  analysis for the combination in next chapter.



**Figure 3.10:** The signal significance for the VBF WW signal as a function of the VBF cuts  $\Delta\eta(j_1, j_2)$  and  $m(jj)$  for signal mass  $700\text{GeV}$  (right). The black outlined bins are those whose values vary from the maximum by less than 5%.

**Table 3.7:** Summary of the selection criteria in the definition of the signal region (SR),  $W$ +jets control region ( $W$  CR) and  $t\bar{t}$  control region ( $t\bar{t}$  CR), in the high-purity (HP) and low-purity (LP) categories.

Selection		SR		W CR		$t\bar{t}$ CR				
		HP	LP	HP	LP	HP	LP			
$W \rightarrow l\nu$	Num of signal leptons	1								
	Num of vetoed leptons	0								
	$E_T^{\text{miss}}$	$> 100\text{GeV}$								
	$p_T(l\nu)$	$> 200\text{GeV}$								
$W/Z \rightarrow J$	Num of large- $R$ jets	$\geq 1$								
	$D_2^{(\beta=1)}$ 50 % WP	pass	fail	pass	fail	pass	fail			
	$D_2^{(\beta=1)}$ 80 % WP	—	pass	—	pass	—	pass			
	$W/Z$ mass 50 % WP	pass	fail	—	—	pass	fail			
	$W/Z$ mass 80 % WP	—	pass	fail	fail	—	pass			
Topology cuts	$p_T(l\nu)/m_{WV}$ $p_T(J)/m_{WV}$	$> 0.3(0.4)$ for VBF (ggF) category								
Top-quark veto	Num of $b$ -tagged jets	0			$\geq 1$					
Multi-jet BG Cleaning Cut	$E_T^{\text{miss}}/p_T(l\nu) > 0.2$	Electron channel only								
Existence of VBF jets		yes (no) for VBF (ggF) category								

### 3.4.5 Boosted Event Selection

In the boosted channel, the most important selection above the others is that at least one large R jet fulfils the  $W/Z$  boson 80% efficiency working points. Then, those events are further categorized into high purity and low purity regions determined by whether the 50% working point is passed. The full selection could be seen in Tab. 3.7

For the leptonically decayed system, the requirement is that exactly one signal is selected with  $E_T^{\text{miss}}$  above  $100\text{GeV}$  to suppress the multijet background. The additional requirements on the system is two topological cuts on kinetic properties:

- (a)  $E_T^{\text{miss}}/p_T(e, \nu) > 0.2$
- (b)  $p_T(e, \nu) > 0.2/m_{WV} > 0.3(0.4)$  for VBF (ggF) category

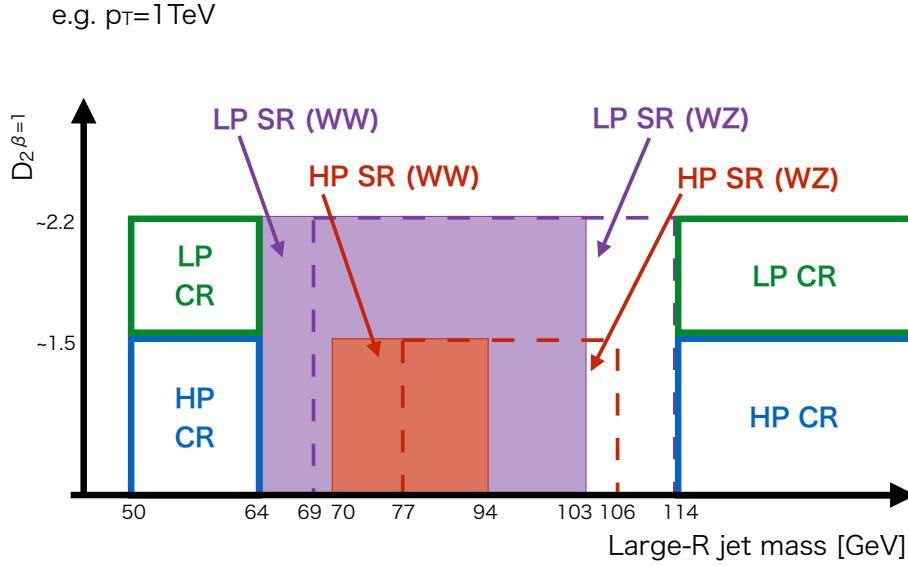
(a) is only for the electron channel to reduced multijet background further in the concern of the jet-faked electrons, while (b) is to lower the SM background contribution for the energy balance between the leptonically and hadronically decayed systems. These criteria are consistent across signal and control regions.

In the side of the hadronically decayed boson is only the large R jet. In addition to the requirement in the last section, the high purity regions (HP) (for both signal and control regions) demand the fat jet boson-tagged at 50% WP, and it is the most sensitive region to signal. For the events with fat jets failing 50% but passing 80% WPs, they went into the low purity region (LP), and the combined sensitivity of HP and LP signal regions could be improved by around 10%. If the fat jet only fails mass cut and pass the  $D_2^{\beta=1}$  of boson tagging, this event would not be discarded but be chosen into W+jet control region instead. Finally,  $p_T$  of the fat jet is also required to be above  $0.3(0.4)m_{WV}$  for the energy balance in VBF(ggF) category. The event categorization of signal and W+jet control regions for both high and low purity categories is illustrated in Fig. 3.11. To reduce the  $t\bar{t}$  background contribution, the subjets associated to the selected large R jet shall not be b-tagged in the W+jet control region and signal region. If any of them or the small R jets ( $R = 0.4$ ) pass the 85% b-tagging WP, the event would go into the top control region.

Fig. 3.12 and Fig. 3.13 are the  $m_{WV}$  distributions for the comparison of signal and background events in high purity and low purity signal regions for electron and muon channels respectively.

### 3.4.6 Resolved Event Selection

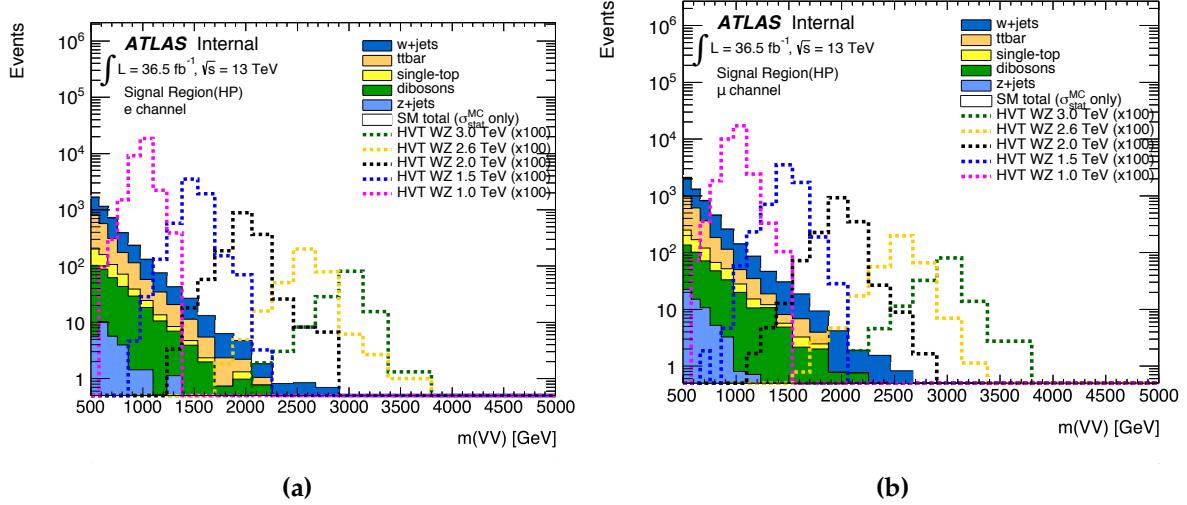
Resolved channel has lower sensitivity than the boosted channel, but it still helps to recover the lost events in the lower energy regime. If the event has no fat jet fulfilling the selection criteria, it went to the resolved category. The full selection for both resolved signal and control regions could be seen in Tab. 3.8



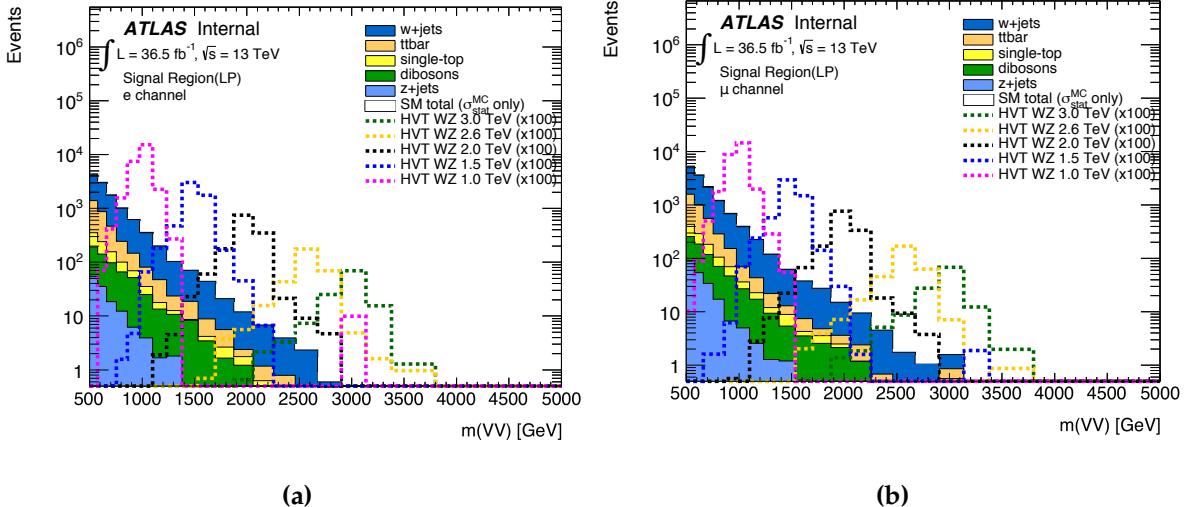
**Figure 3.11:** Definitions of signal region (SR) and  $W+\text{jets}$  control region (WR) for the event with the large-R jet of  $p_T = 1 \text{ TeV}$  based on fat jet boson tagging and number of b-tagged jets

Same to the boosted channel, the system of leptonically decayed  $W$  boson has a signal lepton fulfilling the object definition in last section. However,  $E_T^{\text{miss}}$  cut here is lowered to  $60 \text{ GeV}$  for the less energetic system as compared to the boosted channel. For the energy balance, the ratio between  $p_T(l, \nu)$  and  $m_{WV}$  shall be over 0.3 (0.35) for VBF (ggF) category.

In the hadronic side, the two signal jets are selected after VBF jets, and they are required to have  $p_T$  above  $60 \text{ GeV}$  ( $45 \text{ GeV}$ ) for the leading (subleading) one to suppress the SM background. As they are decayed from  $W$  or  $Z$  boson, their combined mass shall be within the dedicated mass windows, which are  $[66, 94] \text{ GeV}$  for the WW signal region and  $[82, 106] \text{ GeV}$  for the WZ signal region. Because the same top control region is used to make constraint in both of the two signal regions, the mass window is set at  $[66, 106] \text{ GeV}$  as the OR condition of  $W$  and  $Z$  mass windows. For the events with the dijet mass falling into the side band region ( $[0, 66] \text{ GeV}$  or  $[106, 200] \text{ GeV}$ ), they are taken into the  $W+\text{jet}$  control. The energy balance requirement here is the same as the



**Figure 3.12:** The  $m_{WV}$  distributions in the HP signal region for (a) electron and (b) muon channel, with the integrated luminosity of  $36.5\text{fb}^{-1}$ . The HVT WZ signals with  $m = 1.0\text{TeV}, 1.5\text{TeV}, 2.0\text{TeV}, 2.6\text{TeV}$  and  $3.0\text{TeV}$  are overlaid scaled to  $100 \times$  cross section



**Figure 3.13:** The  $m_{WV}$  distribution in the LP signal region for (a) electron and (b) muon channel, with the integrated luminosity of  $36.5\text{fb}^{-1}$ . The HVT WZ signals with  $m = 1.0\text{TeV}, 1.5\text{TeV}, 2.0\text{TeV}, 2.6\text{TeV}$  and  $3.0\text{TeV}$  are overlaid scaled to  $100 \times$  cross section.

leptonic system:  $p_T(jj)/m_{WV} > 0.3(0.35)$  for VBF (ggF) category.

For the optimization of flavour tagging, the existence of b-jets increases the sensitivity in the resolved channel, so the jets could be b-tagged. In WW (WZ) signal region

**Table 3.8:** Summary of the selection criteria of the resolved analysis for the WW and WZ signal regions (SR), W+jets control region (WR) and  $t\bar{t}$  control region (TR).

cuts		WW SR	WZ SR	WR	TR
$W \rightarrow \ell\nu$ selection	Number of signal leptons			1	
	Number of veto leptons			0	
	$E_T^{\text{miss}}$			$> 60\text{GeV}$	
	$p_T(\ell\nu)$			$> 75\text{GeV}$	
$W/Z \rightarrow jj$ selection	Number of small jets			$\geq 2$	
	$p_T(j1)$			$> 60\text{ GeV}$	
	$p_T(j2)$			$> 45\text{ GeV}$	
	$m_{jj}$	[66, 94]GeV	[82, 106]GeV	$< 66\text{GeV}$ or [106, 200]GeV	[66, 106]GeV
Topology cuts	$\Delta\phi(j, \ell)$			$> 1.0$	
	$\Delta\phi(j, E_T^{\text{miss}})$			$> 1.0$	
	$\Delta\phi(j, j)$			$< 1.5$	
	$\Delta\phi(l, E_T^{\text{miss}})$			$< 1.5$	
	$p_T(ev)/m_{WV}$			$> 0.3(0.35)$ for VBF (ggF) category	
	$p_T(jj)/m_{WV}$				
Top veto	Number of $b$ -tagged jets in $W/Z$	$\leq 1$	$\leq 2$	$\leq 1$	$\geq 2$
	Number of other $b$ -tagged jets		0		or $\geq 1$
Existence of VBF jets		yes (no) for VBF (ggF) category			

or W+jet control region, 1 (2)  $b$ -tagged jets can be found in the signal jets, but the presence of more  $b$ -tagged jets in addition to the signal jets would make the event go to the top control region.

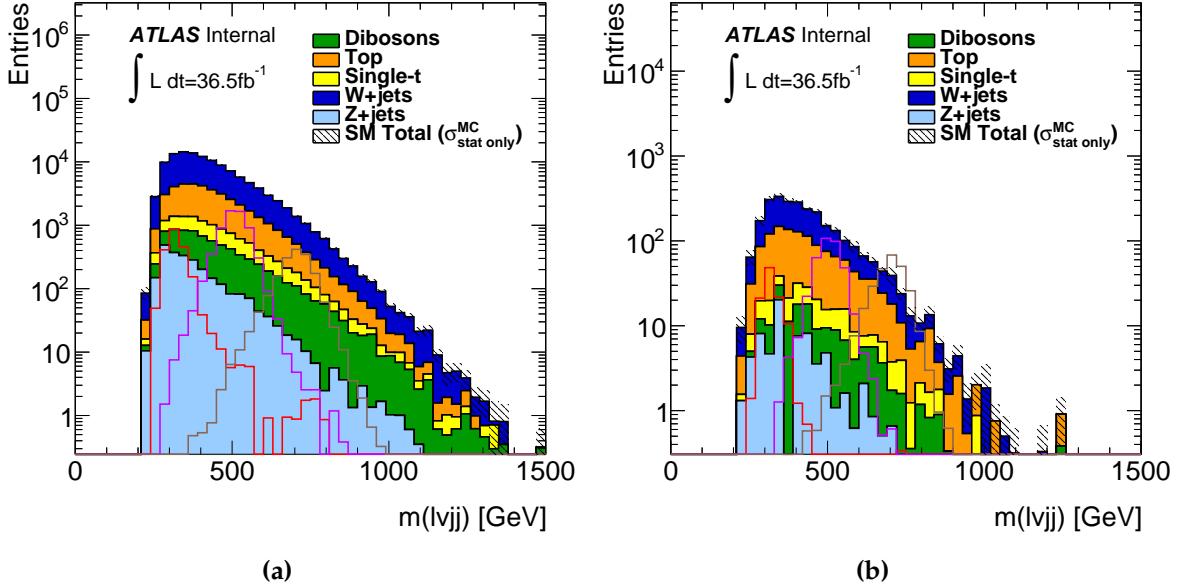
Different from the boosted channel, it has abundant background contribution from multijet background (details in next section), so a series of topological cuts are applied on the signature topology to further suppress it, which are listed below:

- $\Delta\phi(j, l) > 1.0$
- $\Delta\phi(j, E_T^{\text{miss}}) > 1.0$
- $\Delta\phi(j, j) < 1.5$
- $\Delta\phi(l, E_T^{\text{miss}}) < 1.5$

The optimization was studied with dijet ( $jj \rightarrow jj$ ) MC samples.

Fig. 3.14 is the  $m_{WV}$  distributions for comparison of signal and background in re-

solved signal regions for ggF and VBF categories respectively. The signal samples are with lower mass, because resolved channel has better sensitivity to them.



**Figure 3.14:** The  $m_{WV}$  distribution in the resolved signal region for (a) ggF and (b) VBF channel, with the integrated luminosity of  $36.5\text{fb}^{-1}$ . The HVT  $WZ$  signals with  $m = 300\text{GeV}$  (red),  $500\text{GeV}$  (violet) and  $700\text{GeV}$  (blue) are overlaid scaled to  $100 \times$  cross section.

### 3.4.7 Multijet Background Estimation

As discussed above, the SM backgrounds are estimated from Monte-Carlo simulation and constrained in the dedicated control regions. However, multijet processes is poorly modelled due to the lack of understanding to QCD, so simulation is not feasible for this background contribution. Its contribution is from the following sources:

- **Photon Conversion:** When photons or pions interact with the detector material, they decay to soft leptons with similar behaviour to signal ones, which is difficult to recognize. This contribution is mainly into electron, while it could be suppressed by combined muon requirement in muon channel
- **Lepton Misidentification:** Soft charged partons could be blocked at ECAL and leave no signature in HCAL, which is identical to electron signatures. In this case,

they are reconstructed as electrons instead of jets. This source only contributes to electron channel.

- **Heavy Hadron Decay:** The decay products of heavy partons also include leptons. If their decay is close to the primary vertex, the decayed leptons are not distinguishable from the prompt ones. Both electrons and muons have the contribution from this source.

As an alternative, the estimation is performed with fake factor method, a data-driven approach. It is only significant in resolved channel while begin negligible in boosted channel. The details of this method could be found in the ATLAS Run2 VHbb analysis [?, ?].

## Methodology

In the method, to be orthogonal to the signal and control region, fake factors are estimated in the region with only one small R jet called single jet control region where the existence of fat jet passing the selection is not allowed ( $p_T^J > 200\text{GeV}$   $\&\& m^J > 50\text{GeV}$ ) to keep the orthogonality to boosted region. This region is then further divided into two subregions by lepton isolation, as shown in table 3.9. With  $p_T(\mu\nu) < 150\text{GeV}$ , an isolated muon trigger is applied with quality, *ivarmedium*, so the isolation requirement for inverse muon is tightened to remove the bias. As region  $p_T(\mu\nu) > 150\text{GeV}$  is using  $E_T^{\text{miss}}$  trigger, so the isolation bias is not presents. Fake factors in the dedicated bins are defined as:

$$f = \frac{N_{\text{event}}(\text{SingleJetSigLepCR})}{N_{\text{event}}(\text{SingleJetInvLepCR})} \quad (3.21)$$

with the binning in Table 3.10. Fake factors have the dependence on lepton eta (this dependence is for the consideration of detector homogeneity) and  $p_T$ . Additional binning on  $E_T^{\text{miss}}$  is applied in electron channel. For both channels, fake factor is estimated in two different regions with  $p_T(\mu\nu) < 150\text{GeV}$  and  $p_T(\mu\nu) > 150\text{GeV}$  to achieve better precision. Fake factor is shown as a function of lepton  $p_T$  in Figure 3.15 for the region of  $p_T(l\nu) > 150\text{GeV}$ . It could be noticed that fake factor for electron

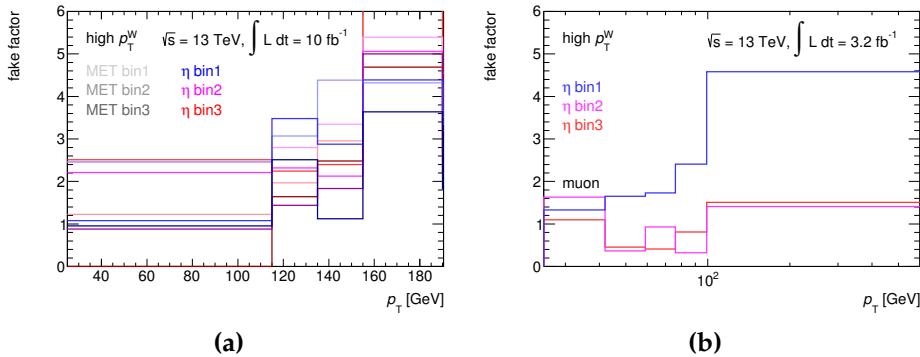
channel is just up to  $p_T = 190\text{GeV}$ . For better accuracy, the fake factor for high  $p_T$  electron is roughly evaluated in  $p_T$  bins only which is shown in Figure 3.16.

**Table 3.9:** Isolation for leptons in the single jet control region

	SingleJetSigLepCR	SingleJetInvLepCR
electron	TightLH	MediumLH (!TightLH)
muon( $p_T(l\nu) > 150\text{GeV}$ )	$Iso_{trk} < 0.06$	$0.06 < Iso_{trk} < 0.15$
muon( $p_T(l\nu) < 150\text{GeV}$ )	$Iso_{trk} < 0.06$	$0.06 < Iso_{trk} < 0.07$

**Table 3.10:** Binning for electrons and muons to evaluate fake factor

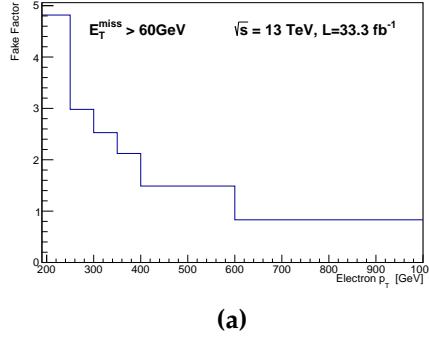
channel	$p_T(\text{GeV})$	$ \eta $	$E_T^{\text{miss}}(\text{GeV})$
electron	27-115	0, 1.37, 1.52, 2.47	0, 60, 75, $\infty$
	115-135		0, 38, 52, $\infty$
	135-155		0, 26, 43, $\infty$
	155-190		0, 25, 45, $\infty$
muon	27, 42, 59, 76, 99, $\infty$	0, 1.05, 1.5, 2.5	N/A



**Figure 3.15:** Fake factors for the corresponding binnings (shown in text) in electron (a) and muon (b) channels

## Electroweak Subtraction

Electroweak interactions ( $t\bar{t}$ ,  $W/Z+\text{jets}$ , diboson and single top) could also contribute to multijet events in addition to the multijet background, so they might be double



(a)

**Figure 3.16:** Fake factors for high  $p_T$  electrons

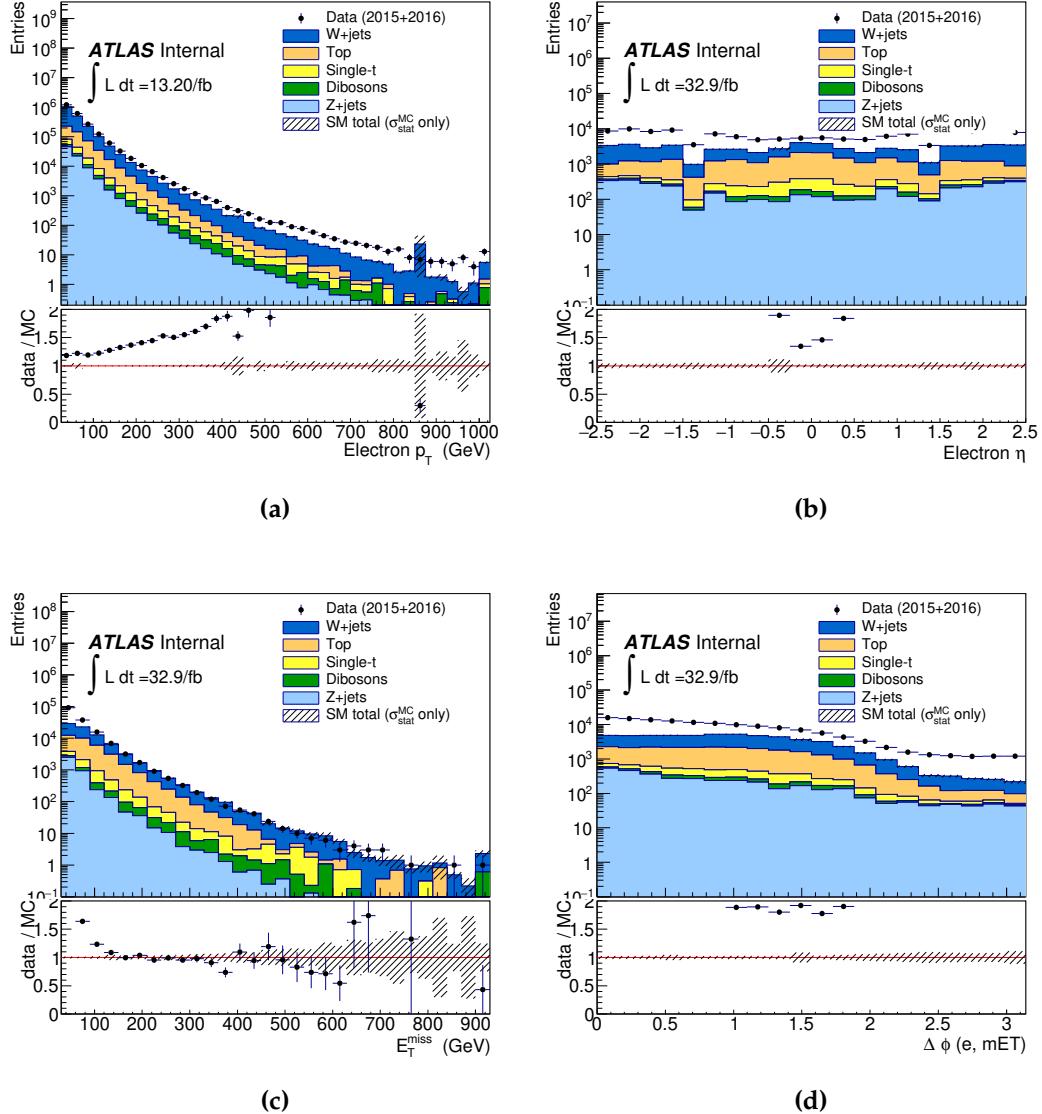
counted from fake factor estimation and Monte Carlo simulation. To avoid this issue, those events are removed by employing fake factor estimation on Monte Carlo samples, which could be expressed as the following equation:

$$N_{events}^{MJ} = N_{events}^{data} - N_{event}^{MC} \quad (3.22)$$

Unfortunately, it is observed that the multijet behavior is not well-modeled in simulation with  $E_T^{miss} > 150\text{GeV}$  for SM backgrounds. This is verified in dijet inversed lepton control region with lepton isolation inversed and at least two jets with  $P_T > 20\text{GeV}$ . Figure 3.17 and Figure 3.18 show the discrepancy between data and electroweak interactions which are contributed from multijet events. It is anticipated that multijet contribution shall be almost 0 in high  $E_T^{miss}$  region in the dijet control region, but the discrepancy is still observed. In this case, the electroweak subtraction is applied with a scale factor derive from the ratio of events from data and simulation in the bin  $150\text{GeV} < E_T^{miss} < 250\text{GeV}$  defined as:

$$f = \frac{N_{event}(data)}{N_{event}(MC)} \quad (3.23)$$

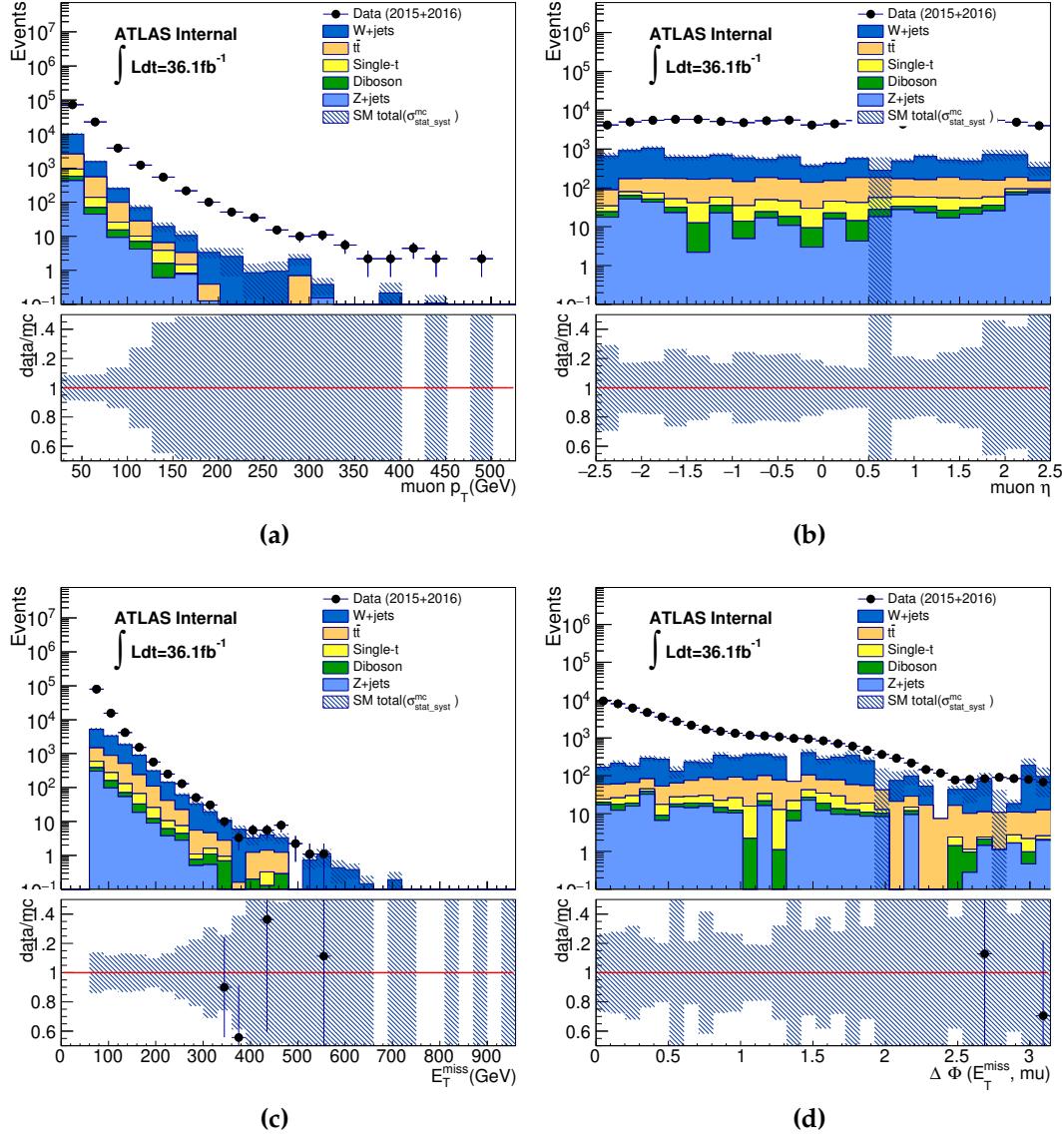
It is applied as an additional correction on fake factor for events with  $E_T^{miss} > 150\text{GeV}$  from simulation. The electroweak subtraction factors for electron and muon channels are shown in Table 3.11.



**Figure 3.17:** The distribution of lepton  $p_T, \eta, E_T^{\text{miss}}$  and  $\Delta\phi(e, E_T^{\text{miss}})$  in dijet fake control region with inversed lepton for electron channel. The inconsistency is thought to be comprised of multijet events without applying electroweak subtraction.

**Table 3.11:** Electroweak subtraction factor for electron and muon channels

channels	electron	muon
EW subtraction factor	1.36	1.49

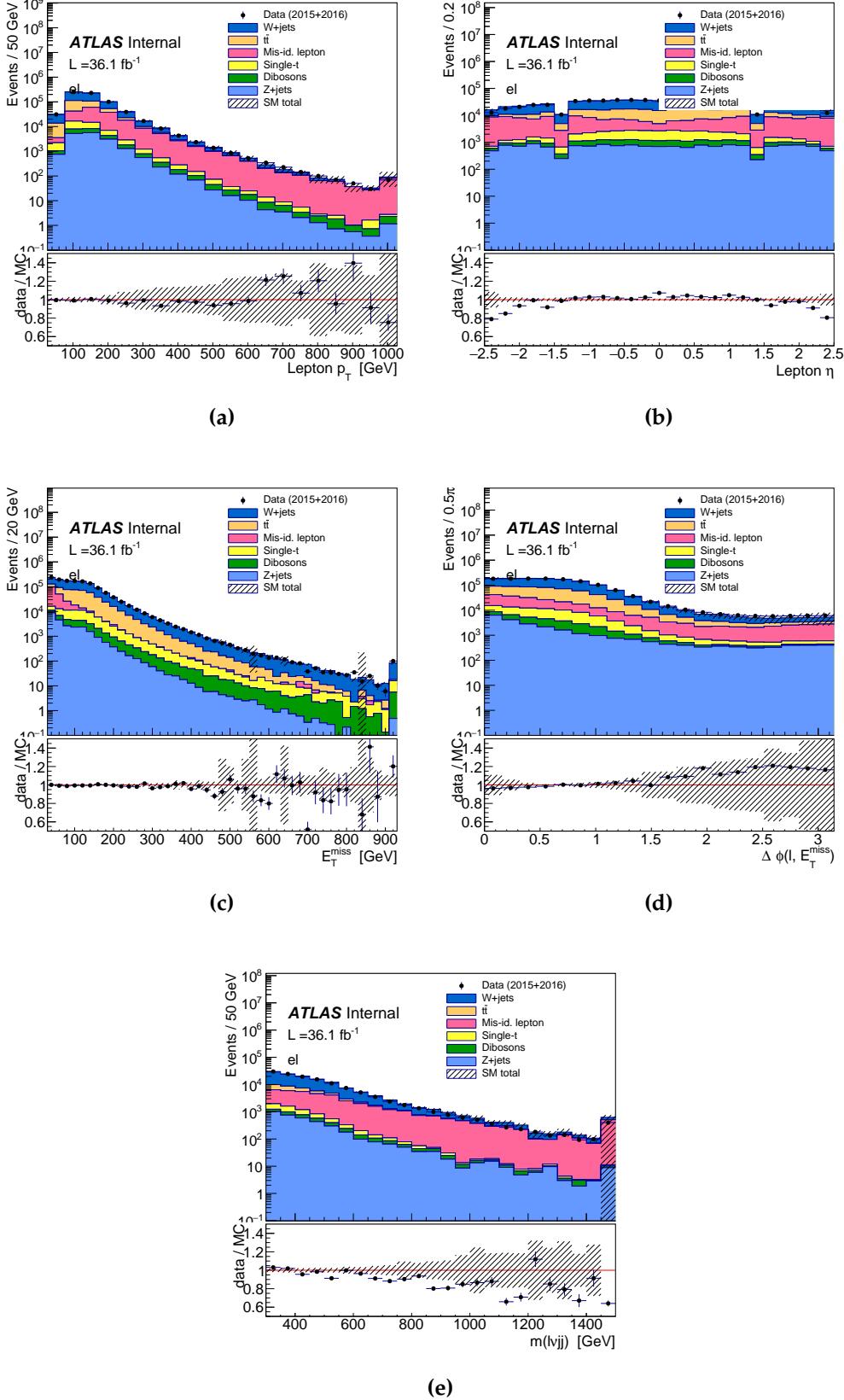


**Figure 3.18:** The distribution of lepton  $p_T, \eta, E_T^{miss}$  and  $\Delta\phi(\mu, E_T^{miss})$  in dijet fake control region with inversed lepton for muon channel. The inconsistency is thought to be comprised of multijet events without applying electroweak subtraction.

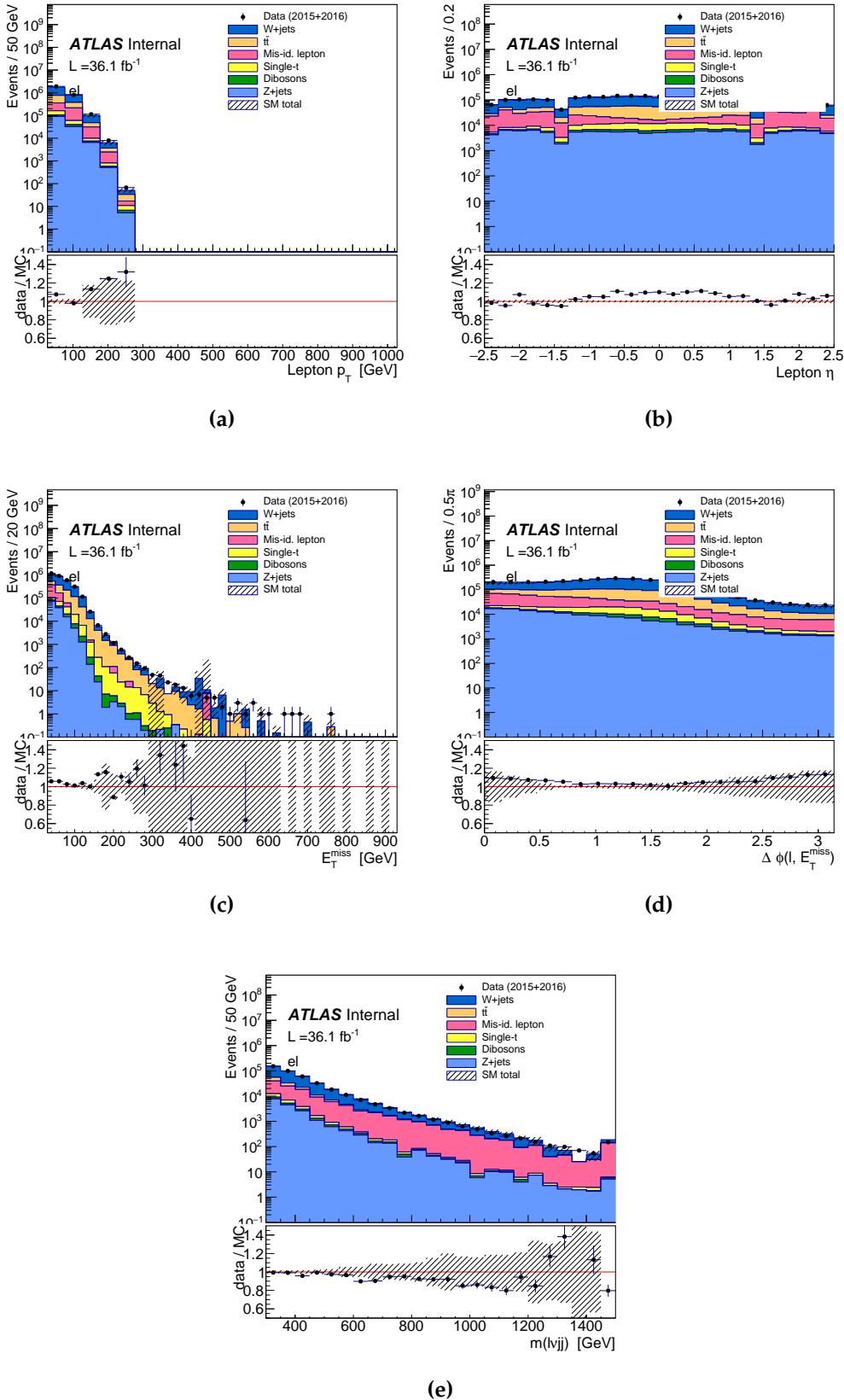
## Validation

The method is validated in the dedicated validation region. The definition is similar to the signal region with looser cut to enrich the multijet events. It requires at least two resolved jets ( $p_T^{leading} > 60\text{GeV}$ ,  $p_T^{subleading} > 45\text{GeV}$ ),  $30\text{GeV} < E_T^{miss} < 50\text{GeV}$ , exactly one isolated lepton and the resolved triggers passed for electron and muon channels respectively. This definition is slightly overlapped with signal and control

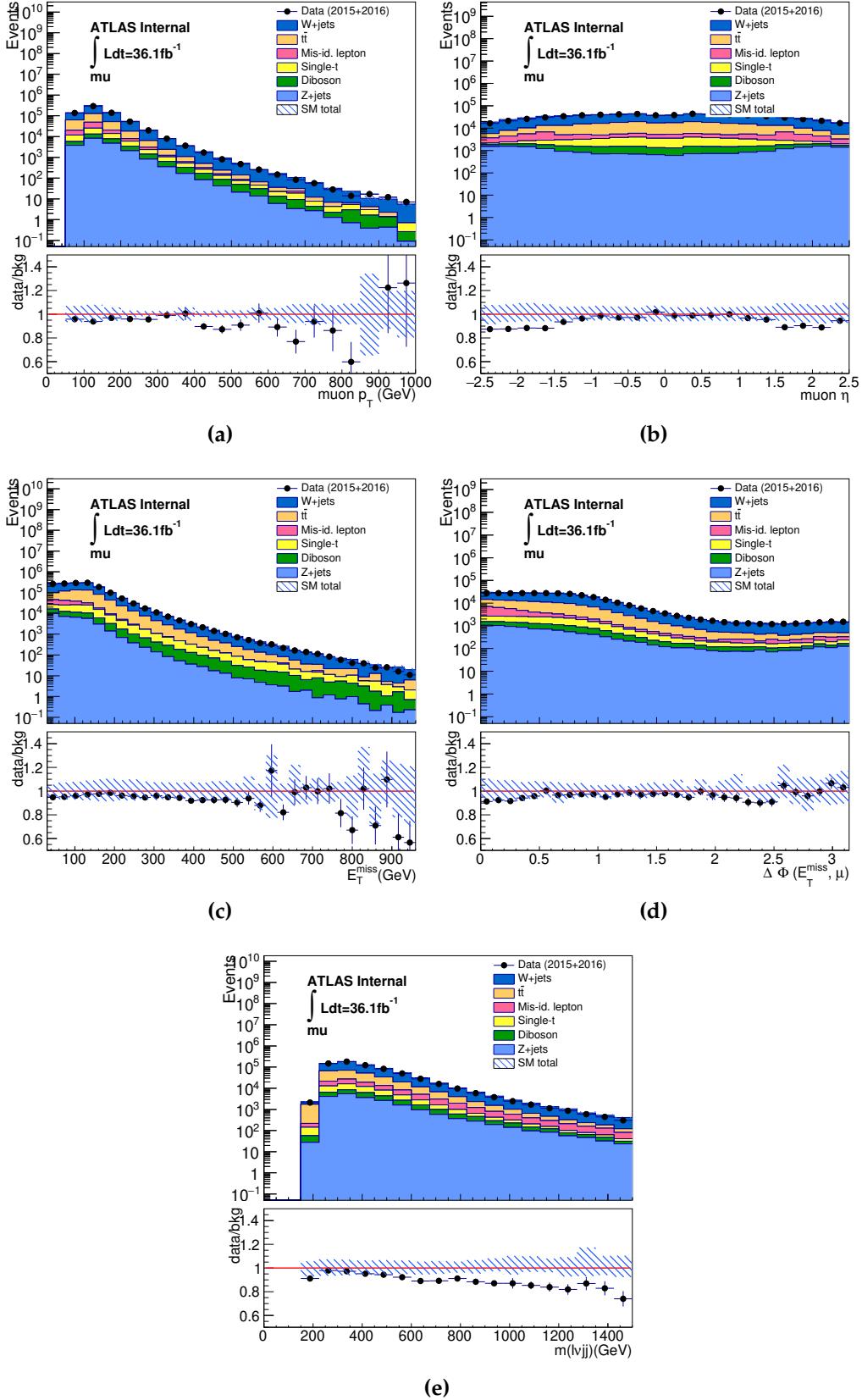
regions, but the upper cut on  $E_T^{miss}$  suppress the signal contribution. As the fake factors were derived from two bins of  $p_T(l\nu)$ , the validation is performed on  $p_T(l\nu) < 150GeV$  and  $p_T(l\nu) > 150GeV$ . The results are presented in Figures 3.19 - 3.22 with multijet background estimated using fake factor method. In general, data agrees well with backgrounds with tolerable inconsistency within statistic uncertainties. The disagreement in the region of  $p_T(l\nu) > 150GeV$  is supposed to be due to the low statistics for fake factor estimation in single jet control region, but it should not have great impact in final interpretation, as multijet events would just account for around 10% of the whole background. The related systematic uncertainty would be discussed in Chapter 4.



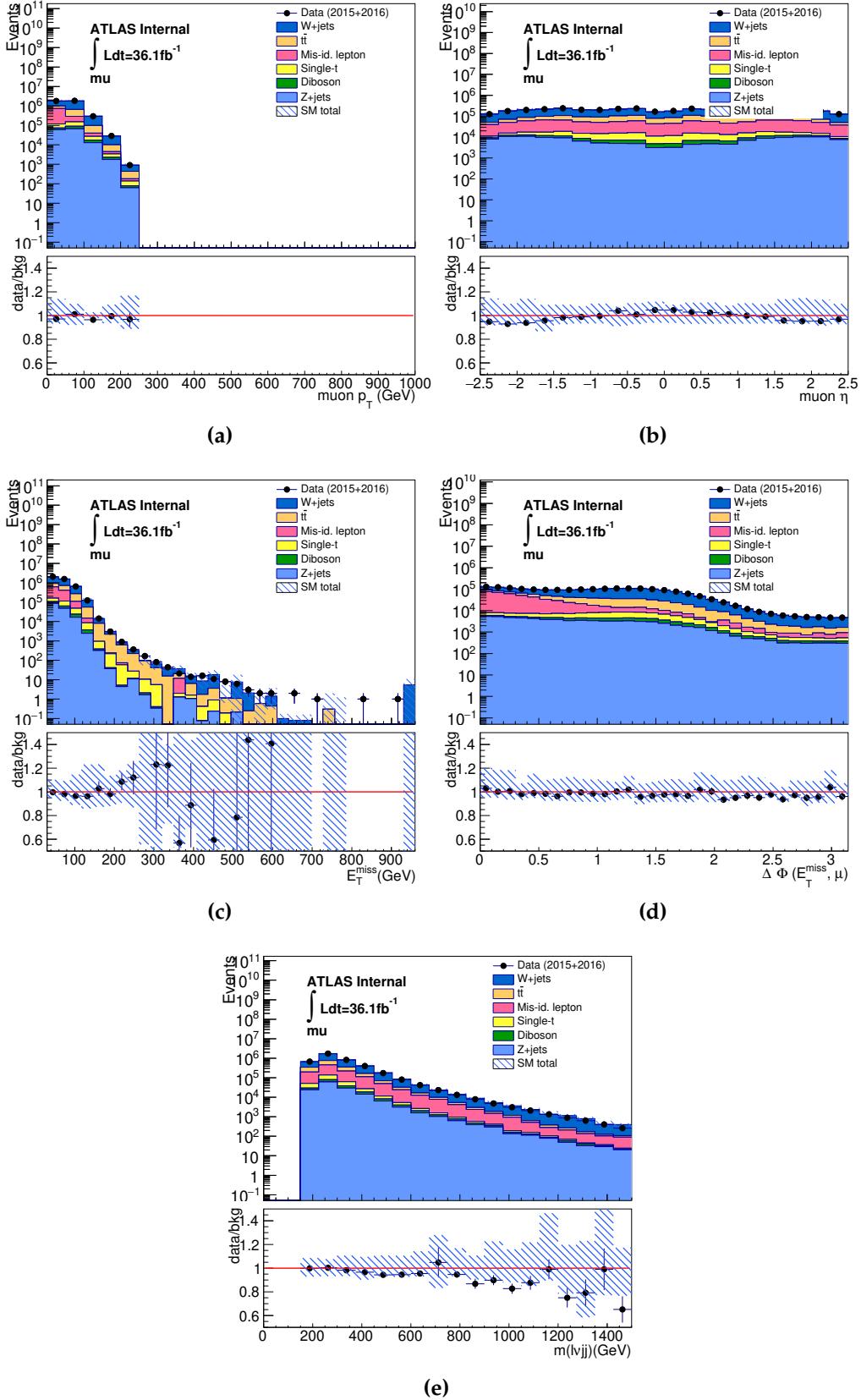
**Figure 3.19:** The distribution of lepton  $p_T, \eta, E_T^{\text{miss}}, \Delta\phi(l, E_T^{\text{miss}}), m_{WV}$  in validation region with  $p_T(l\nu) > 150 \text{ GeV}$  in electron channel with multijet background



**Figure 3.20:** The distribution of lepton  $p_T, \eta, E_T^{\text{miss}}, \Delta\phi(e, E_T^{\text{miss}}), m_{WV}$  and BDT in validation region with  $p_T(l\nu) < 150\text{GeV}$  in electron channel with multijet background



**Figure 3.21:** The distribution of lepton  $p_T, \eta, E_T^{\text{miss}}, \Delta\phi(\mu, E_T^{\text{miss}}), m_{WV}$  and BDT in validation region with  $p_T(l\nu) > 150\text{GeV}$  in muon channel with multijet background

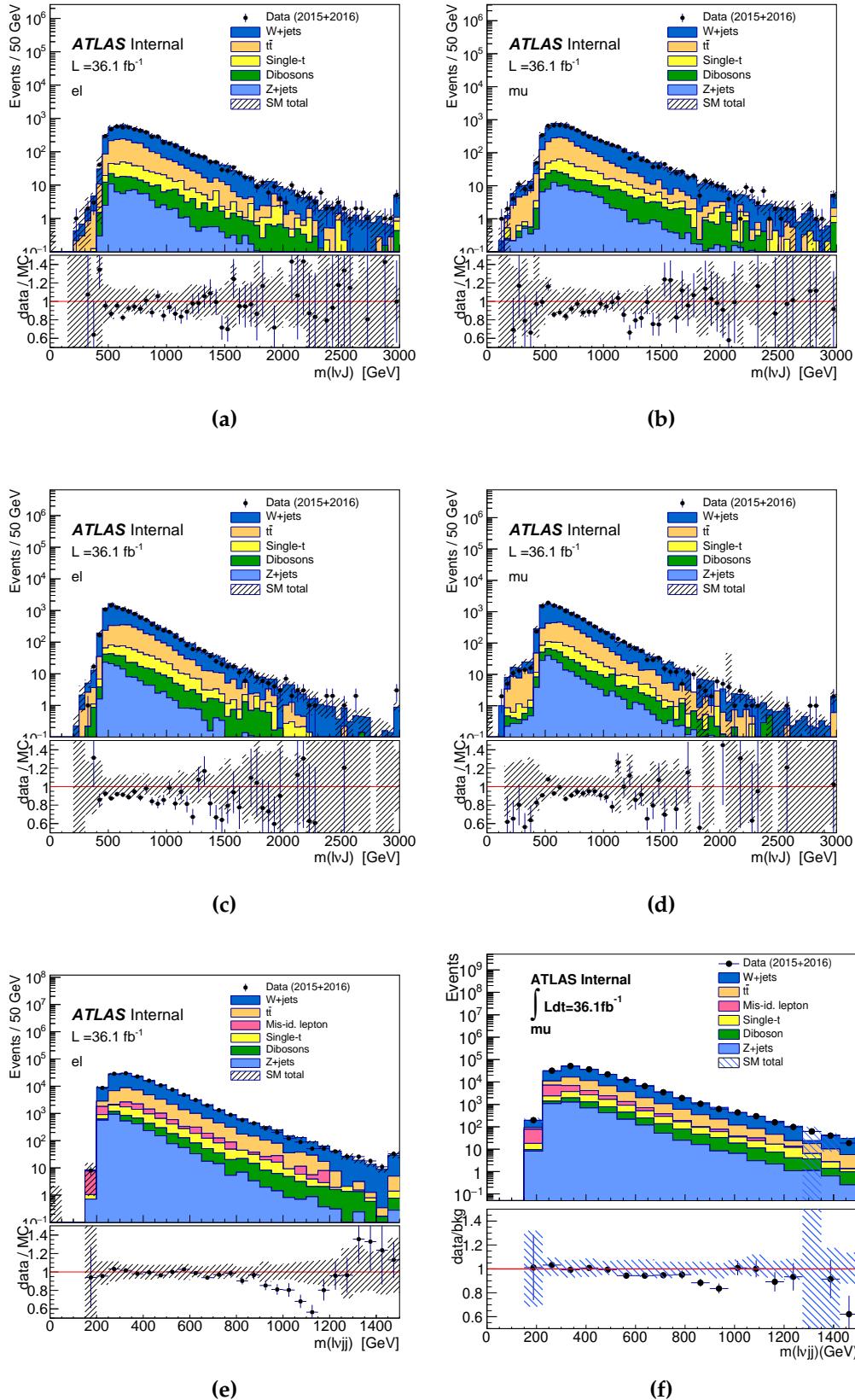


**Figure 3.22:** The distribution of lepton  $p_T, \eta, E_T^{\text{miss}}, \Delta\phi(\mu, E_T^{\text{miss}}), m_{WV}$  and BDT in validation region with  $p_T(l\nu) < 150\text{GeV}$  in muon channel with multijet background

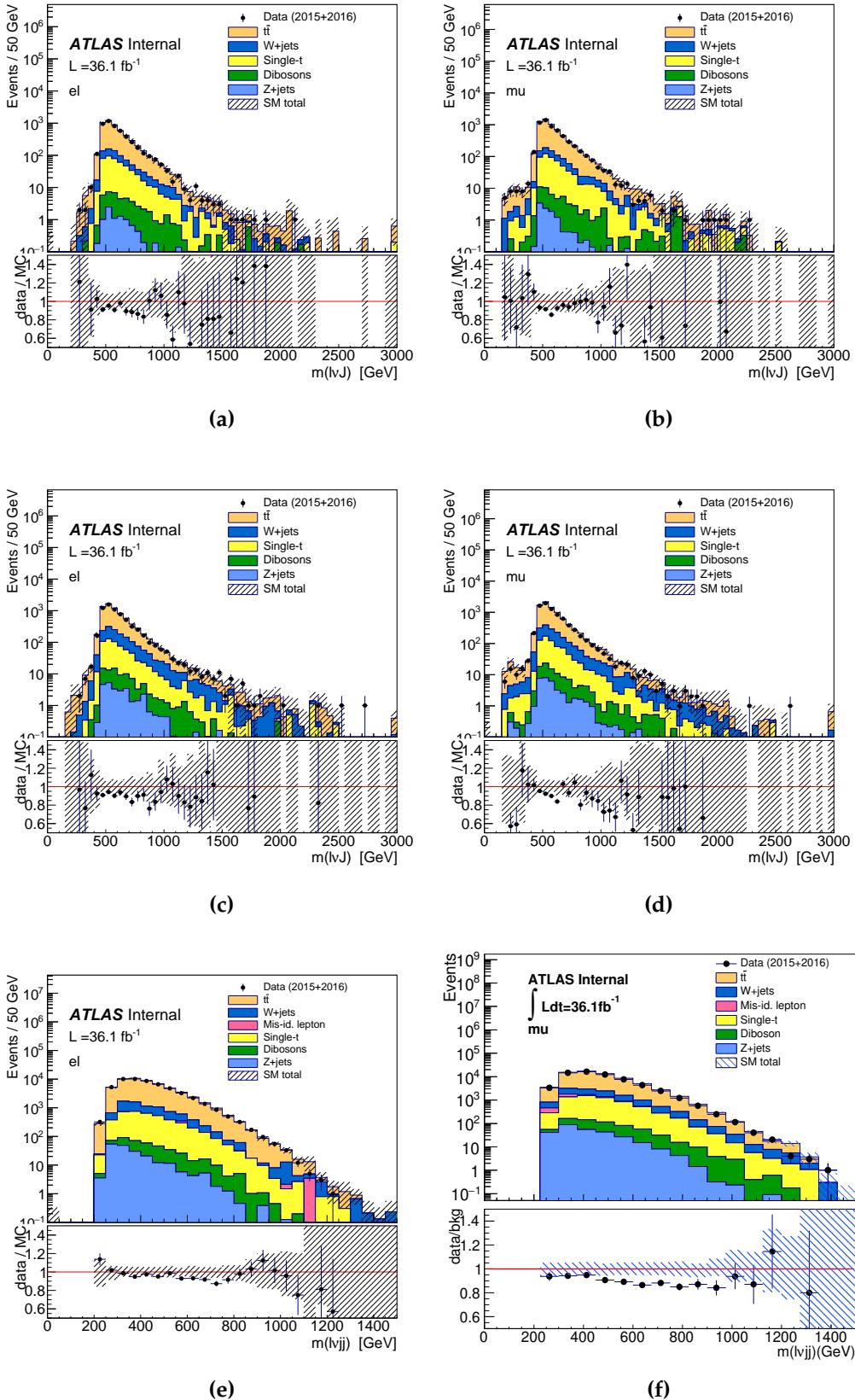
### 3.5 Data Background Comparison

To verify the modelling of background estimation, the comparison in top and W+jet control regions are performed for both VBF and ggF categories. The consistency is not perfect as expectation, and the fitting in the control regions is on the purpose to recover it, which will be discussed in next chapter. The other issue in the background simulation is that a slope in the ratio of data over background is observed in  $m_{jj}^{VBF}$  in resolved VBF category for V+jet samples from Sherpa generator. In this analysis, it is also taken as one contribution to the mismodelling of simulation.

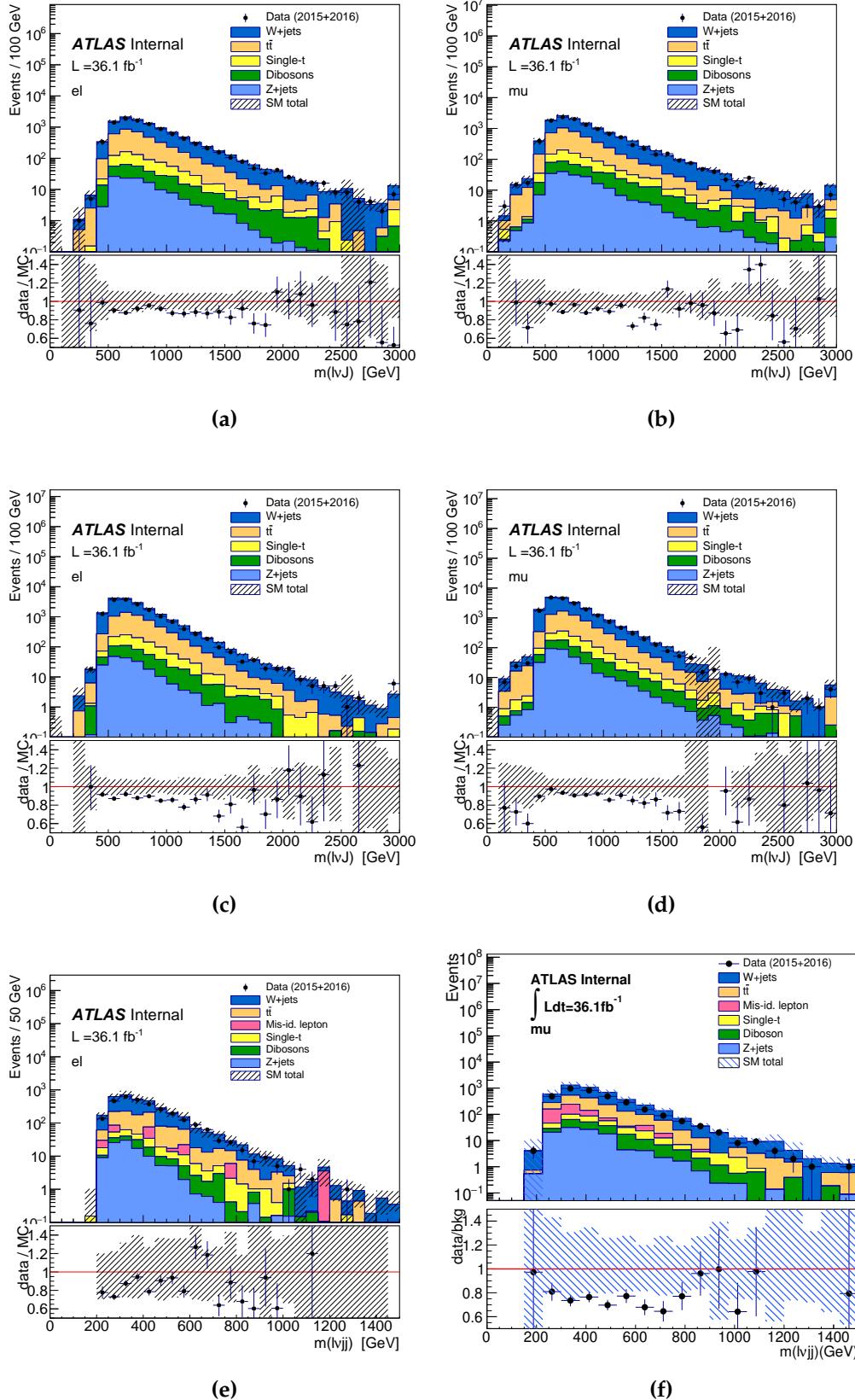
Fig. 3.23 and Fig. 3.24 are the comparison plots for  $m_{WV}$  in ggF category, while Fig. 3.25 to Fig. 3.27 are for VBF category. The comparison of  $m^{VBF}(j,j)$  could be found in Fig. 3.26 and Fig. 3.28 to examine the VBF modelling.



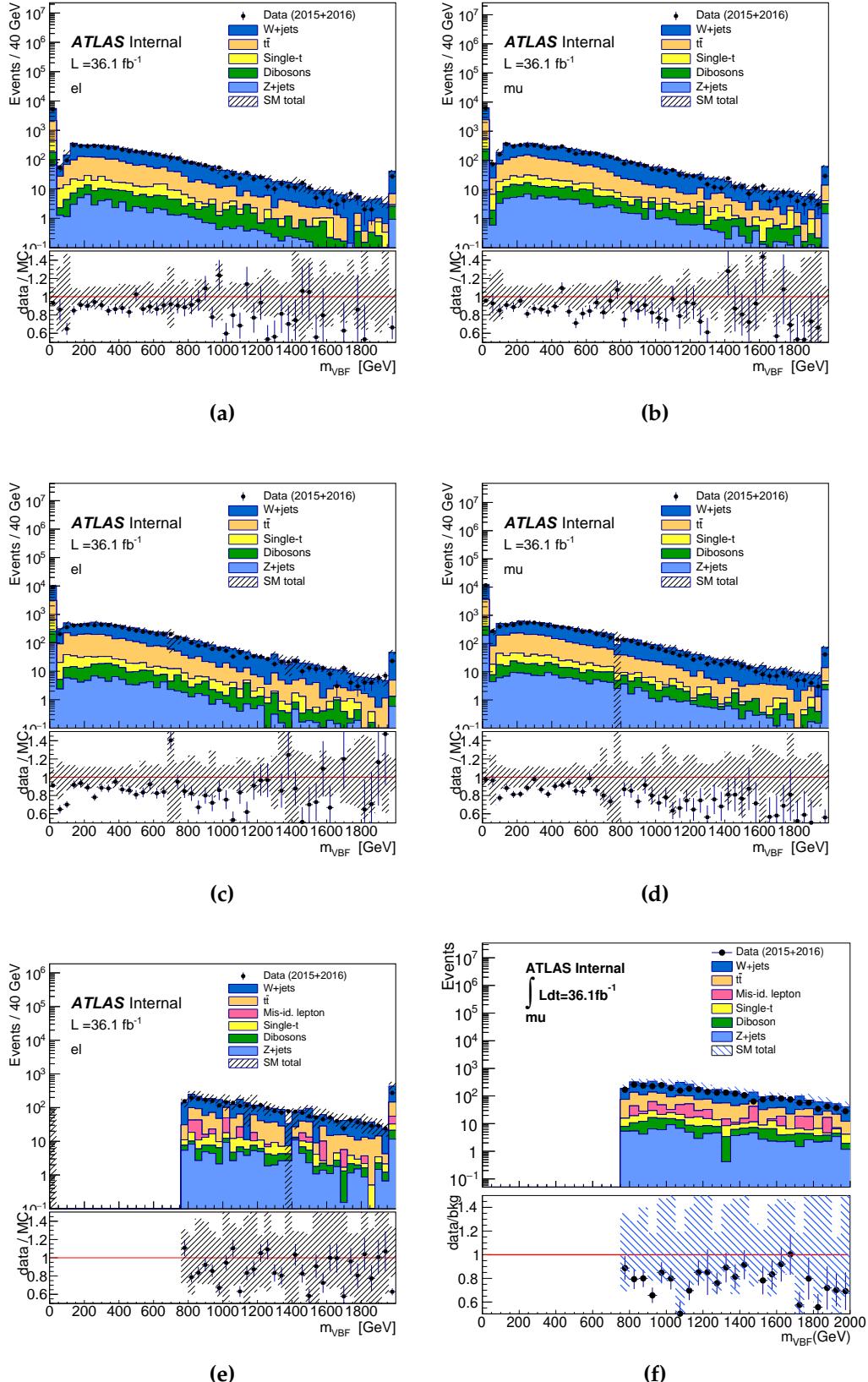
**Figure 3.23:** The distribution of  $m_{WV}$  in ggF high purity (top), low purity (middle), and resolved (bottom) W+jet control region for electron (left) and muon (right) channels respectively



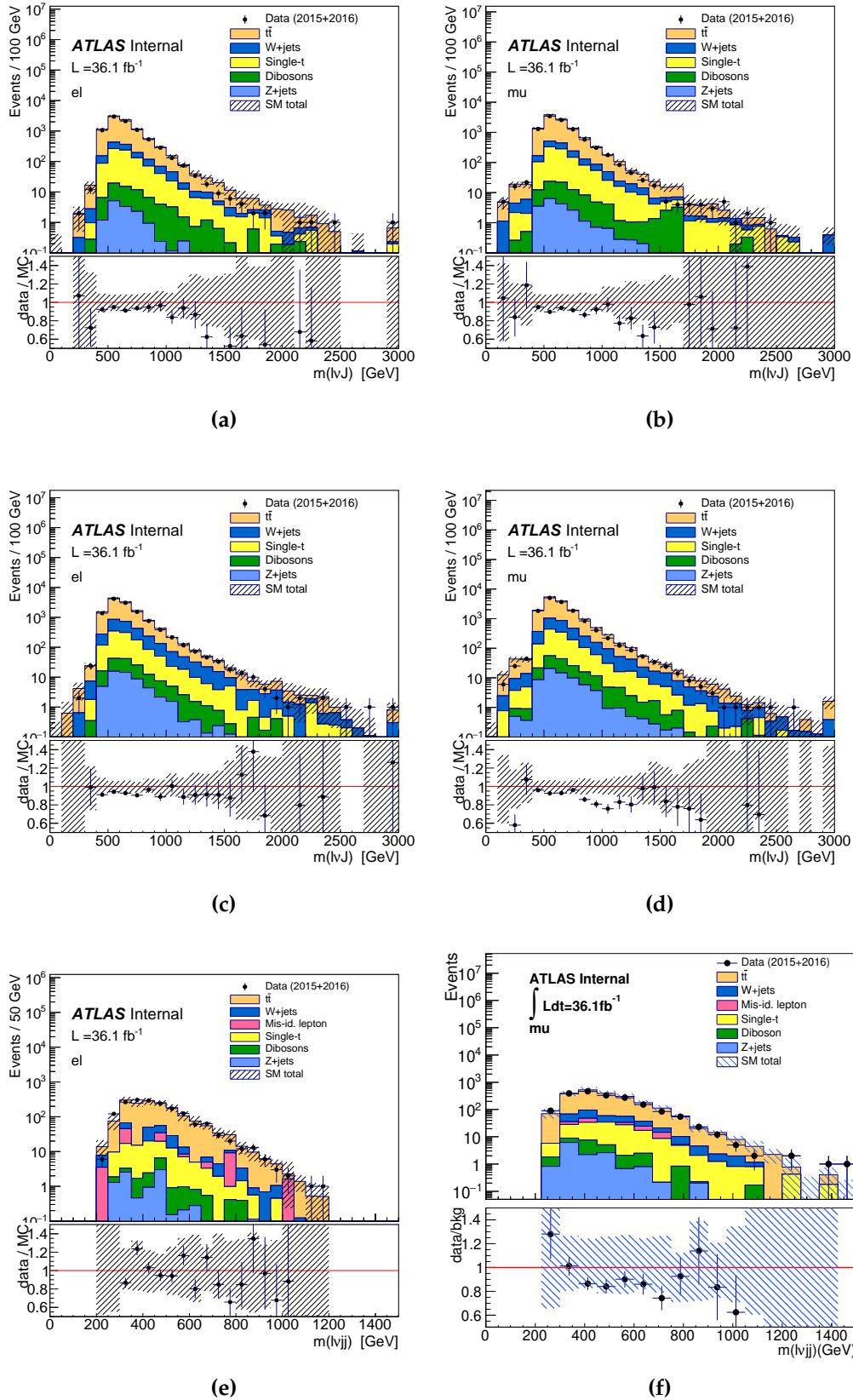
**Figure 3.24:** The distribution of  $m_{WV}$  in ggF high purity (top), low purity (middle), and resolved (bottom) top control region for electron (left) and muon (right) channels respectively



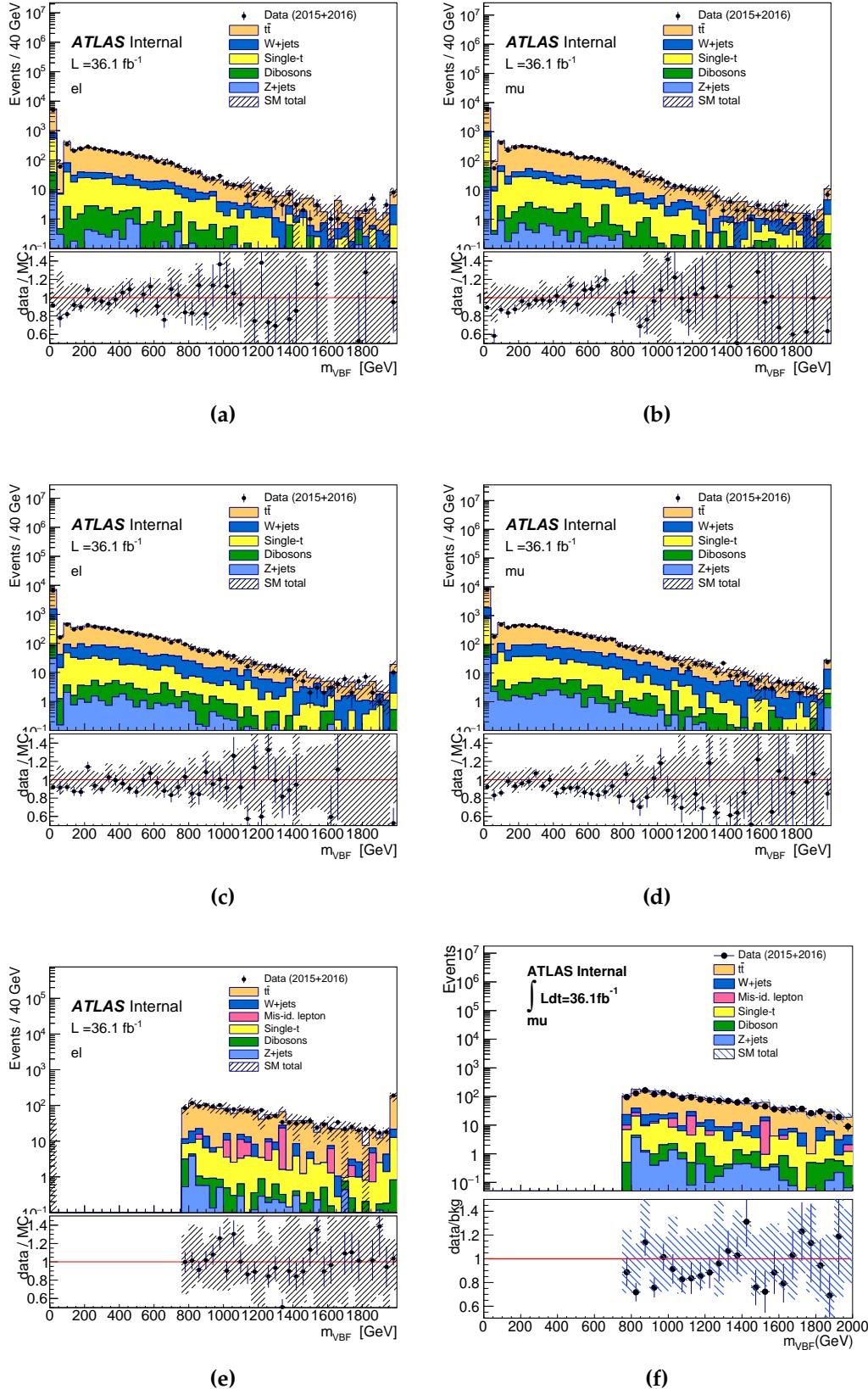
**Figure 3.25:** The distribution of  $m_{WV}$  in VBF high purity (top), low purity (middle), and resolved (bottom) W+jet control region for electron (left) and muon (right) channels respectively



**Figure 3.26:** The distribution of  $m_{jj}^{VBF}$  in VBF high purity (top), low purity (middle), and resolved (bottom) W+jet control region for electron (left) and muon (right) channels respectively



**Figure 3.27:** The distribution of  $m_{WV}$  in VBF high purity (top), low purity (middle), and resolved (bottom) top control region for electron (left) and muon (right) channels respectively



**Figure 3.28:** The distribution of  $m_{jj}^{VBF}$  in VBF high purity (top), low purity (middle), and resolved (bottom) top control region for electron (left) and muon (right) channels respectively.

# Chapter 4

## Interpretation for Resonance Analysis

After having the  $m_{WV}$  distributions from both control and signal regions, the statistic interpretation is applied to verify whether any signal signature is captured in this analysis. It will be going through the following steps:

- Variation on Histograms: the systematic uncertainties contributed from the background modelling and experiments are applied in the analysis, and each of them vary the  $m_{WV}$  distribution in histograms. They are taken as the inputs with the nominal histogram for the background fitting in the next step.
- Control Region Fitting: a binned maximum-likelihood fitting is performed in control and signal region histograms simultaneously to rescale the backgrounds and signal for a proper agreement to the data. The scale factors are then taken as the ratio of post-fit to pre-fit histograms in each bin.
- Signal Verification: the signal interpretation is through the CLs method by quantifying the agreement between data and background in signal regions after signal fitting. The result will be presented as the exclusion on the mass regions at 95% confidence level or the discovery with a corresponding “p-value”.

The details of each step will be discussed in the following sections with the results from this analysis.

## 4.1 Systematic Uncertainties

No measurement and theoretical estimation could be 100% accurate, and the uncertainties could propagate to the  $m_{WV}$  histograms. In this case, a bump in data might be due to the uncertainty fluctuation but mistaken as a signal. To prevent this mistake, both systematic and statistic uncertainties are brought into the consideration for the ground fitting and signal interpretation.

The following are the systematic uncertainties considered in this analysis and how they are taken into the  $m_{WV}$  histograms.

- **Luminosity Measurement:** the given luminosity of the dataset collected in 2015 and 2016 is accompanied by the uncertainty of 2.1%. It is applied in the histograms from simulations by scaling up and down the total yield of each bin by 2.1%
- **Selection and Reconstruction Efficiency:** the object reconstruction and selection efficiency of physical objects are not consistent between data and simulation like the trigger efficiency shown in Subsec. 3.4.1. The uncertainty of this source is induced by the uncertainties in the variables used in tag and probe method. To estimate the impact, the tag and probe criteria are tightened and loosened for scale factor re-estimation, and they replace the nominal scale factors to obtain the new histograms. This uncertainty comes from the efficiencies of trigger, lepton isolation, lepton identification, jet b-tagging, fat jet boson-tagging, and all physical object reconstruction.
- **Energy Scale and Resolution:** the energy measurement is based on the pulse shapes from the calorimeter cells, but it is not precise enough due to different responses of layers or varied granularity of the calorimeter. The uncertainty estimation of this source for electrons and muons are via the Z boson mass reconstruction in dedicated analysis as a function of  $p_T$ . In the case of jets, they are estimated via the comparison of the truth and reconstructed  $E_T$  from dijet simulation samples. It also has the impact on  $E_T^{miss}$  reconstruction, and the variation on jet energy scale is the dominant contribution for  $E_T^{miss}$  uncertainty.

The variation from the uncertainty is applied as the variations on object  $E_T$  in the analysis to get the new  $m_{WV}$  histograms.

- **Simulation Modelling:** The tuning and modelling parameters are different for generators and showering models due to the varied preference of theoretical approximation. To take this variation into the uncertainty contribution, simulated samples are regenerated with another simulation sets (a different generator or tuning parameters), and the same events selections is applied. The new histogram is then obtained after the normalization to the nominal sample. This is contributed from  $W + jet$ ,  $t\bar{t}$  and signal simulation (for signal, the tuning is to consider the events for which there are the jets from the initial or final state radiation (ISR and FSR)). As other backgrounds have minor contribution, the effect is taken negligible.
- **Multijet Background Modelling:** multijet modelling is sensitive to the lepton isolation criteria and the jet topology. To estimate the uncertainty of this contribution, the fake factors were re-evaluated with loosened and tightened isolations on leptons as well as in the single b-jet control region, and the new fake factors are applied to get the new multijet  $m_{WV}$  distribution.

## 4.2 Likelihood Construction & Fitting

A simultaneous fitting is conducted to adjust the background and signal to agree well with the data in the  $m_{WV}$  histogram which is in the binnings of boosted and resolved channels:

$$m_{WV}^{Boosted} = 500, 575, 660, 755, 860, 975, 1100, 1235, 1380, 1535, 1700, \\ 1875, 2060, 2255, 2460, 2675, 2900, 3135, 3380, 3800, 6000 \quad (4.1)$$

$$m_{WV}^{Resolved} = 300, 360, 420, 500, 575, 660, 755, 860, 975, 1100, 1500, 2000 \quad (4.2)$$

For the VBF category, the bins with higher  $m_{WV}$  have the statistics too low for the MC samples, so there is only one bin for  $m_{WV} > 1535 \text{ GeV}$ ( $1100 \text{ GeV}$ ) for the boosted (resolved) region. Then, a maximum likelihood method is performed for the fitting which could be presented in the full form as:

$$\mathcal{L}(\mu, \theta) = \prod_k \left\{ \prod_{i=1}^{N_{\text{bins},k}^{\text{SR}}} P(N_{ki}^{\text{SR}} | \mu s_{ki}^{\text{SR}} + \mu_{t\bar{t},k} b_{t\bar{t},ki}^{\text{SR}} + \mu_{W,k} b_{W,ki}^{\text{SR}} + b_{\text{others},i}^{\text{SR}}) \times \right. \\ \prod_{l=1}^{N_{\text{bins},k}^{\text{TR}}} P(N_{kl}^{\text{TR}} | \mu s_{kl}^{\text{TR}} + \mu_{t\bar{t},k} b_{t\bar{t},kl}^{\text{TR}} + \mu_{W,k} b_{W,kl}^{\text{TR}} + b_{\text{others},m}^{\text{TR}}) \times \\ \left. \prod_{m=1}^{N_{\text{bins},k}^{\text{WR}}} P(N_{km}^{\text{WR}} | \mu s_{km}^{\text{WR}} + \mu_{t\bar{t},k} b_{t\bar{t},km}^{\text{WR}} + \mu_{W,k} b_{W,km}^{\text{WR}} + b_{\text{others},m}^{\text{WR}}) \right\} \\ \times \prod_{j=1}^{N_\theta} \text{Nuis}(\theta_j), \quad (4.3)$$

where  $P(a|b)$  is the Poisson probability distribution function (p.d.f.) to observe “a” number of events (data) when “b” number of events is expected from theory (background and signal estimation) in each bin. To properly normalize the background,  $\mu s$  are the most important parameter in the formula as floating parameters to rescale the event numbers in each region for background estimation , and they shared between control and signal regions (simultaneously). The  $\mu$  to rescale the signal events is also called signal strength which is the primary parameter of interest in the statistical interpretation. The  $k$  index in this formula corresponding to the event categories: ggF merged HP, ggF merged LP, ggF resolved, VBF merged HP, VBF merged LP, and VBF resolved regions, and their likelihoods are constructed independent from each other.

### Nuisance Parameters

The last term in Eq. 4.3 is to take in the consideration of uncertainties mentioned in the last section. They are called “nuisance parameters” in the scope of statistics, as they only have the impact on the shape of likelihood which is of the second interest. Our primary parameter of interest (POI) is  $\mu$ , the scale factor for signal events.

There are three types of nuisance parameters based on their impact on the distribution of  $m_{WV}$ . The following are the treatments to them mostly with the constraint of a Gaussian distribution in this analysis, although the other p.d.f. options are also available.

- Statistical Uncertainty: with the limited event numbers of background estimation, the statistical uncertainties in each bin are taken as extra nuisance parameters. A light Beeston-Barlow method is applied which introduces a new scale factor,  $\theta$ , on each bin with the constraint of a Gaussian distribution with the default value as 1. These nuisance parameters are then contributed to the likelihood in this expression:

$$Nuis(\theta) = \prod_i \exp \left[ \frac{(\theta_i - 1)^2}{2\sigma_i^2} \right] \quad (4.4)$$

$\theta$  is the ratio of the scaled event number to the unscaled (raw) event number in the prediction. The likelihood is then further constrained by the Gaussian distribution of  $\theta$  which has the width of  $\sigma$  from quadratic sum of all the background contributions.  $i$  is still the index of each bin.

- Overall Normalization: this type of nuisance parameters has the equally sided uncertainties, and they just scale up and down the total yields of histograms without changing the shape of distribution. They are contributed by uncertainties from the scaling factors and luminosity measurement. The treatment is simply taking a Gaussian distribution as the constraint if it is needed. It can be presented in the likelihood as:

$$Nuis(\theta) = \exp \left[ \frac{(\theta - N)^2}{2\sigma^2} \right] \quad (4.5)$$

In this expression, the Gaussian distribution has the mean of observed event number with the width of observed uncertainty for luminosity. In the case of uncertainties for scale factors, Tab. 4.1 is the summary for constraints applied on different background contributions. For  $t\bar{t}$  and W+jets backgrounds, no constraint is set, and the deviation of  $\theta$  from 1 for the scale factors is always taken as  $1\sigma$ .

**Table 4.1:** The constraints on scaling factors for SM backgrounds

Background	Constraint	Upper Limit ( $\sigma_+$ )	Lower Limit ( $\sigma_-$ )
W+jets	Free	2	0
$t\bar{t}$	Free	2	0
single top	Gaussian	1.11	0.89
WW+WZ	Gaussian	1.3	0.7
Z+jets	Gaussian	1.11	0.89

- Shape Related Uncertainty: for the uncertainties which are not equally sided ( $\sigma_+ \neq \sigma_-$ ), a procedure called “morphing” is applied which could be presented as:

$$n = \begin{cases} n_0 + \theta(\sigma_+ - n_0) & \theta > 0 \\ n_0 + \theta(n_0 - \sigma_-) & \theta < 0 \end{cases} \quad (4.6)$$

Here,  $n$  is the scaled event number, while  $n_0$  is the raw event number. Then, scaled factor is constrained by  $\theta$  which is under a Gaussian distribution constraint ( $G(\mu, \sigma) = G(0, 1)$ ).

## Quality of Fitting

To find the maximum of likelihood in Eq. 4.3, the logarithmic form,  $\log \mathcal{L}$ , is used. The extreme value is then found when:

$$-\frac{\partial}{\partial \mu} \log \mathcal{L} = 0 \quad (4.7)$$

However, the phase space of the likelihood is complex constructed with multiple dimensions of the scale factors, so the MINUIT2 method with Hessian matrix is performed under the framework of RooFit. The maximized likelihood is denoted as:  $\mathcal{L}(\hat{\mu}, \hat{\theta})$

To verify the quality of the process of fitting with the likelihood equation, two properties of the results are verified:

- **Pull** The pull is defined as the deviation of nuisance parameters from the expected mean number:

$$pull = \frac{\hat{\theta} - \theta_0}{\sigma_\theta} \quad (4.8)$$

with  $\theta_0$  as the mean of  $\theta$ , while the uncertainty of nuisance parameters are taken from the likelihood phase space. The pull result is verified by the comparison to “Asimov data” which took the expected event number as the observed data. (so the Asimov data has the pull as 0.) The proper fitting should have all the pulls within the  $1\sigma$  variation with the reasonable uncertainty, or that indicates a huge discrepancy between the background estimation and the observed data.

- **Nuisance Parameter Correlation** The phase space of likelihood is constructed under the assumption that all the nuisance parameters are decorrelated, but it still needs to be verified. The correlation matrix is then used for this verification which has the elements defined as:

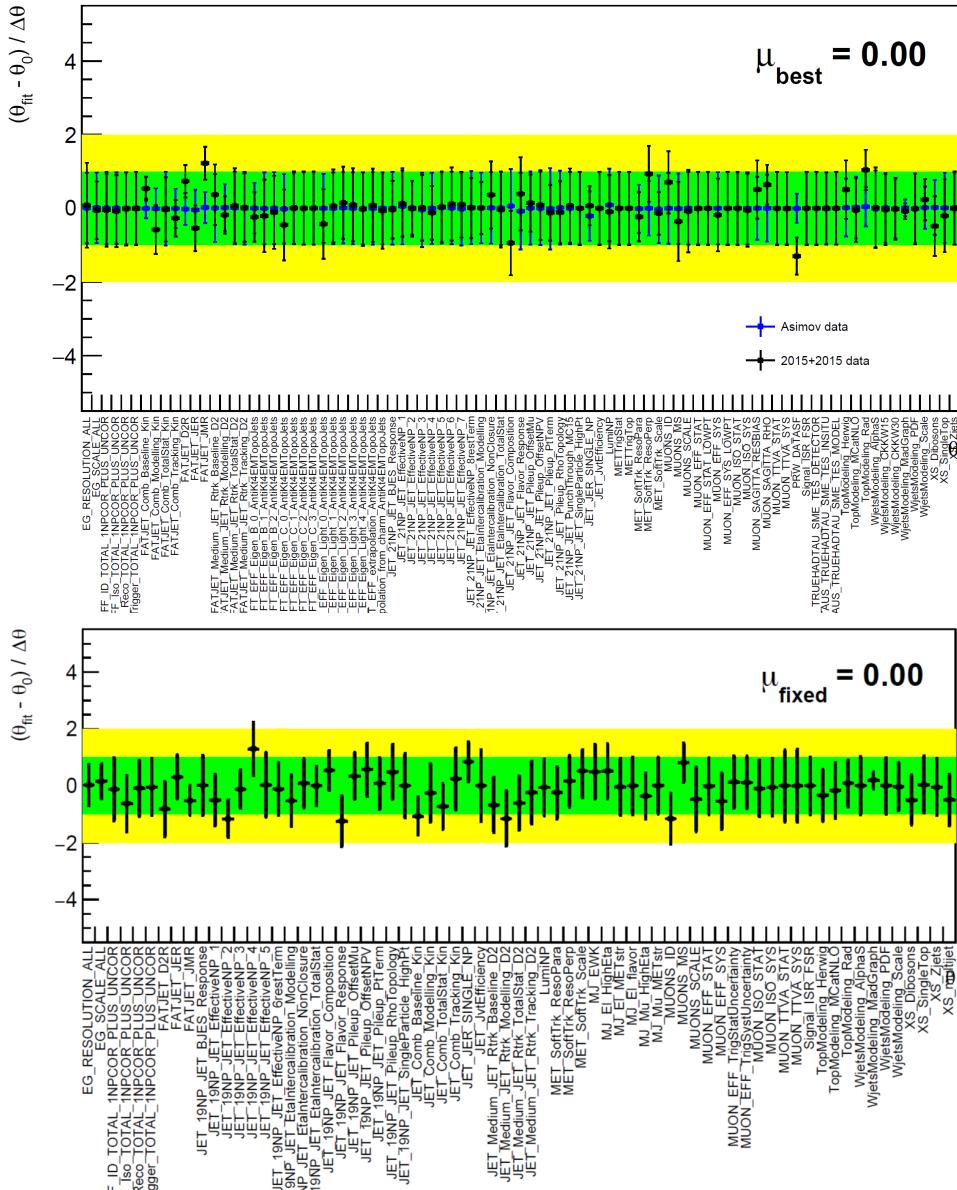
$$Cov(\theta_i, \theta_j) = \frac{\partial^2 \log(\mathcal{L})}{\partial \theta_i \partial \theta_j} \Big|_{\theta=\hat{\theta}} \quad (4.9)$$

This variable,  $Cov(\theta_i, \theta_j)$ , should be close to 0 if  $i \neq j$ .

The pulls are with the signal of ggF 2000 GeV and 500 GeV W' bosons for boosted and resolved categories which are presented in Fig. 4.1 with signal strength (signal scale factor) as 0 in control regions. Fig. 4.2 is the correlation matrix of the nuisance parameters applied in the ggF HP boosted region. **Combination**

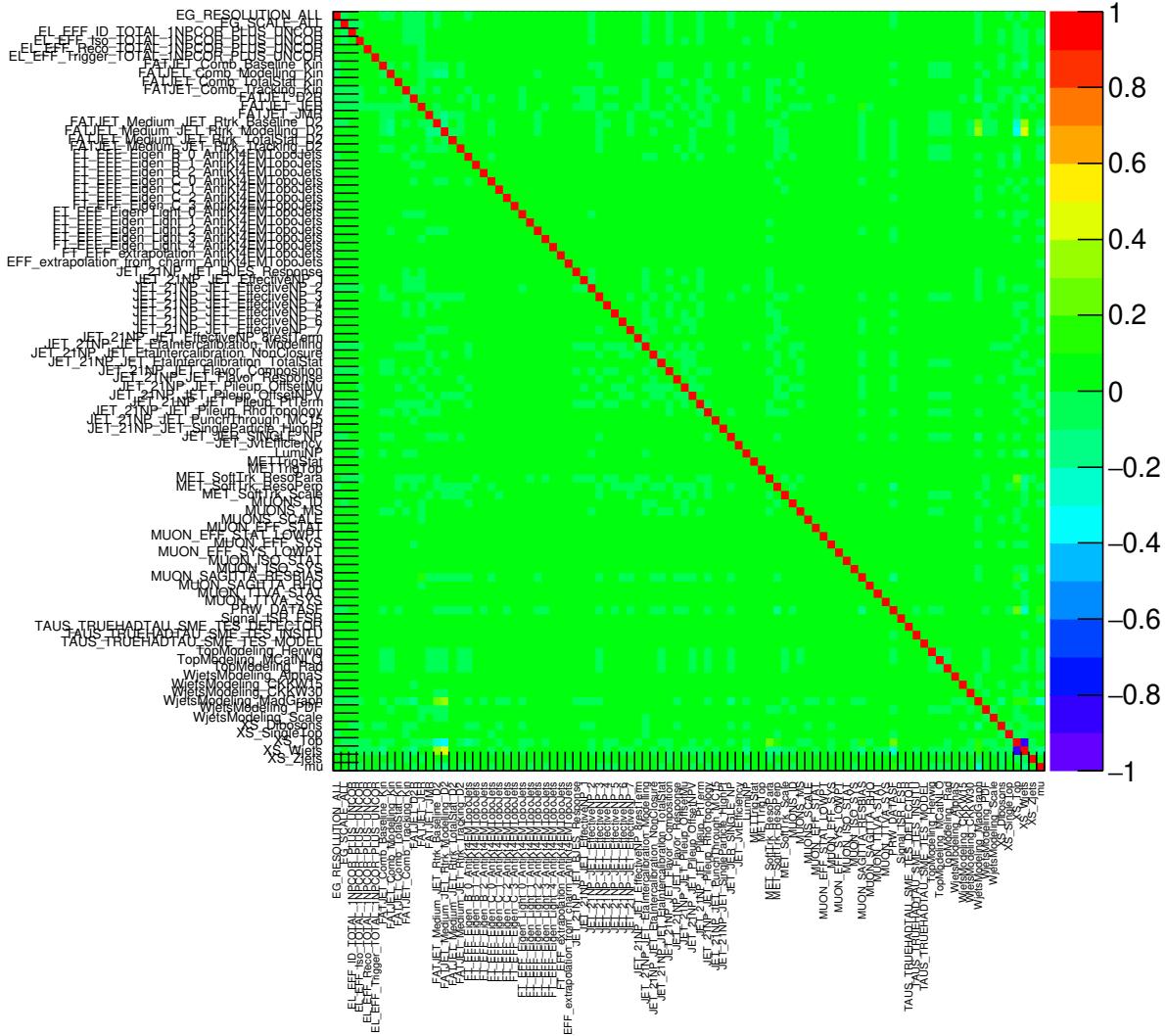
This analysis contains several categories (merged and resolved, or VBF and ggF production), and a combination of them could help to increase the sensitivity to set a more stringent limit by the decrease of distribution width in the test statistics p.d.f. (this will be discussed in the next section). The combination procedure is to simply multiply the likelihoods constructed from different event categories:

$$\mathcal{L} = \prod_{i=1}^{N_{categories}} \mathcal{L}_i(\mu, \theta_i) \quad (4.10)$$



**Figure 4.1:** The pulls for the fitting with input signal of ggF 2 TeV (up) and 500 GeV (down) W' bosons for the boosted and resolved categories respectively.

The signal strength,  $\mu$ , would be common across the likelihoods. For the nuisance parameter terms, if they are from the same source like the uncertainty in object energy measurements, they are also the same among the event categories. In this case, those nuisance parameters are “correlated”. For the other case, when the nuisance parameters are from an independent source which is not considered in the other



**Figure 4.2:** The correlation matrix of boosted high purity region with the ggF event selection

category like the multijet uncertainties in the resolved category, it would only make the constraint on the likelihood of this category, and they are called “decorrelated”.

### 4.3 Conclusion and Result

After the fitting, the agreement between data and background+signal expectation event numbers should be verified to test whether the “hypothesis” of existence or exclusion of signal is correct. The final interpretation is conducted in two ways: the

exclusion for setting limits and the significance of a discovery.

### Methodology for a Discovery (p-Value)

This is a counting analysis for which the property we want to measure is to see where a signal bump could be spotted in the diboson mass spectrum, so a profile likelihood with the likelihood built in the last session is formulated (for the case of a precision measurement, the “Neyman–Pearson lemma” is preferred in the format of  $\lambda = \mathcal{L}(H_1)/\mathcal{L}(H_0)$ ) to simplify the phase space to verify the varied signal strength:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (4.11)$$

where  $\mathcal{L}(\hat{\mu}, \hat{\theta})$  is the maximized likelihood with  $\hat{\mu}$  and  $\hat{\theta}$ , while  $\mathcal{L}(\hat{\mu}, \hat{\theta})$  is the maximized likelihood with a specific  $\mu$  by giving  $\hat{\theta}$ . The test statistics is then constructed as  $-2 \ln \lambda(\mu)$ . Following by this, a test statistics is built which is given the form:

$$q_0 = \begin{cases} -2 \ln \lambda(0) & 0 \leq \hat{\mu} \\ 0 & \hat{\mu} < 0 \end{cases} \quad (4.12)$$

For the second case of  $\hat{\mu} < 0$ , this is not to reject the background only hypothesis. However, the derivation of a p.d.f. for the test statistics is computationally expensive, so an asymptotic approach is applied. The first step is to apply the Wald approximation, and the test statistics could be simplified to:

$$-2 \ln(\lambda(\mu)) = \left( \frac{\mu - \hat{\mu}}{\sigma} \right)^2 + \mathcal{O}(1/N) \quad (4.13)$$

with  $\sigma$  taken as the uncertainty in the likelihood phase space along the  $\mu$  direction and  $N$  is the observed event number. However, to evaluate  $\sigma$  is computationally expensive, so, in this analysis, the Asimov data is used. With Eq. 4.13,  $\sigma$  could be evaluated as:

$$\sigma^2 = \frac{\mu - \hat{\mu}}{-2 \ln(\lambda(\mu))} \quad (4.14)$$

With an enough event number, the last term in Eq. 4.13 is negligible. From Wilks theorem, if a hypothesized  $\mu'$  is true, the probability of measuring a specific  $\hat{\mu}$  should follow a Gaussian distribution:

$$\hat{\mu} \sim Gaus(\mu', \sigma) \quad (4.15)$$

Then, the probability distribution of the test statistics would be in a “chi-square distribution” which is written as  $f(q_\mu | \mu')$  with the non-central parameter as:

$$\Lambda = \left( \frac{\mu - \mu'}{\sigma} \right)^2 \quad (4.16)$$

For a discovery with the test statistics in Eq. 4.12,  $\mu'$  is set to 0, and a “p-value” is then defined as:

$$p_0 = \int_{q_{0,obs}}^{\infty} f(q_0 | \mu' = 0) dq_0 \quad (4.17)$$

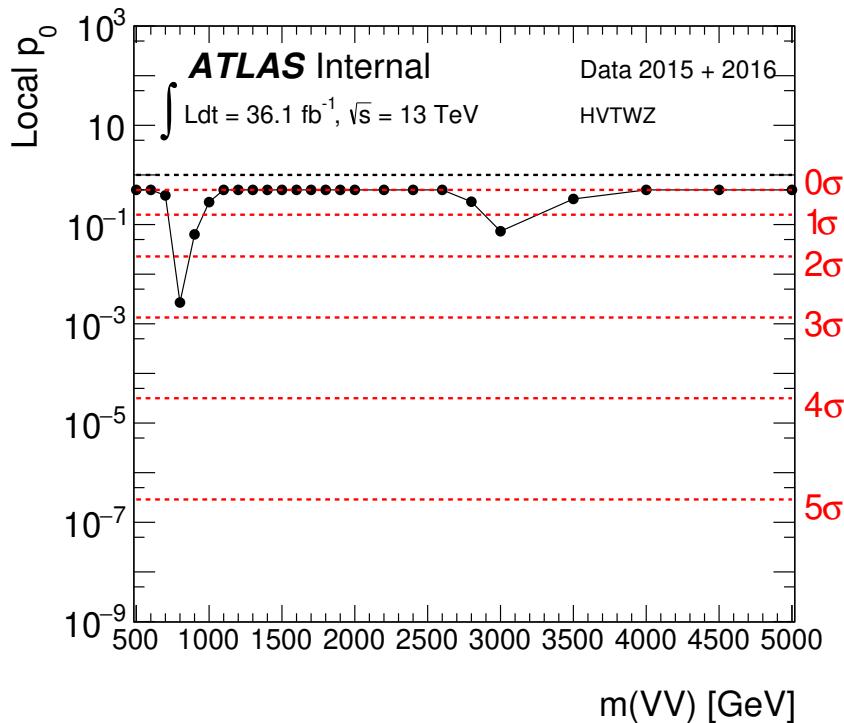
where  $q_{0,obs}$  is taken at the  $\mu$  value which gives the observed event yield. This is indicating the possibility that the null hypothesis ( $\mu' = 0$ ) is wrong, and it shows great disagreement to data. p-value would also be interpreted into the discovery significance:

$$Z = \Phi^{-1}(1 - p_0) \quad (4.18)$$

where  $\Phi^{-1}$  is the quantile for inverse cumulative distribution of a standard Gaussian. Fig. 4.3 shows the p-value and discovery significance for the ggF HVT signal combined with both resolved and boosted regions. The best significance is given at 800 GeV for less than  $3\sigma$ . In particle physics, the discovery of a new particle could only be claimed with an excess of  $5\sigma$  which is tight to avoid the so-called “type-I error” defined as making a false discovery.

### Methodology for an Exclusion (Confidence Interval at 95% Confidence Level)

Without a significant result ( $Z < 3\sigma$ ), an exclusion limit is then set to conclude that a



**Figure 4.3:** The observed p-value and significance for the W' boson from the ggF production with the combined data of both resolved and merge channels

specific range of theoretical hypotheses (i.e. new particles of varied mass range) has no signal which is within the analysis sensitivity (i.e. the particle production cross-section is significant to be measured).

In the case of an exclusion, an alternative test statistics is formulated as:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \lambda(\mu) & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \\ -2 \ln \lambda(0) & \hat{\mu} < 0 \end{cases} \quad (4.19)$$

For the three cases in the expression, the bottom one is to keep  $\mu$  positive to have physical meaning, when  $\hat{\mu}$  is smaller than 0. For the other two cases, it is just to have the  $\mu$  hypothesis at one side for  $\mu > \hat{\mu}$  which is to set the exclusion limit.

Then, the asymptotic approach is applied again. Under this case,  $\tilde{q}_\mu^*$  is chosen with the

Asimov data to make:

$$p_\mu = \int_{\tilde{q}_\mu^*}^{\infty} f(\tilde{q}_\mu | \mu = 0) d\tilde{q}_\mu = 0.05 \quad (4.20)$$

This is meaning that if the signal exists with a specific signal strength,  $\mu^*$ , the null hypothesis would be rejected at 95% confidence level (CL). Followed by that,  $\mu^*$  is taken as the median value for the new p.d.f.,  $f(\tilde{q}_\mu | \mu = \mu^*)$ , and also the expected upper limit of sensitivity to measure the signal. Then, the observed sensitivity is estimated to be the  $\mu$  in this new p.d.f. corresponding to the observed event yield.

The final result is then interpreted by converting the evaluated  $\mu$  into the production cross-section and the decay branch ratio:

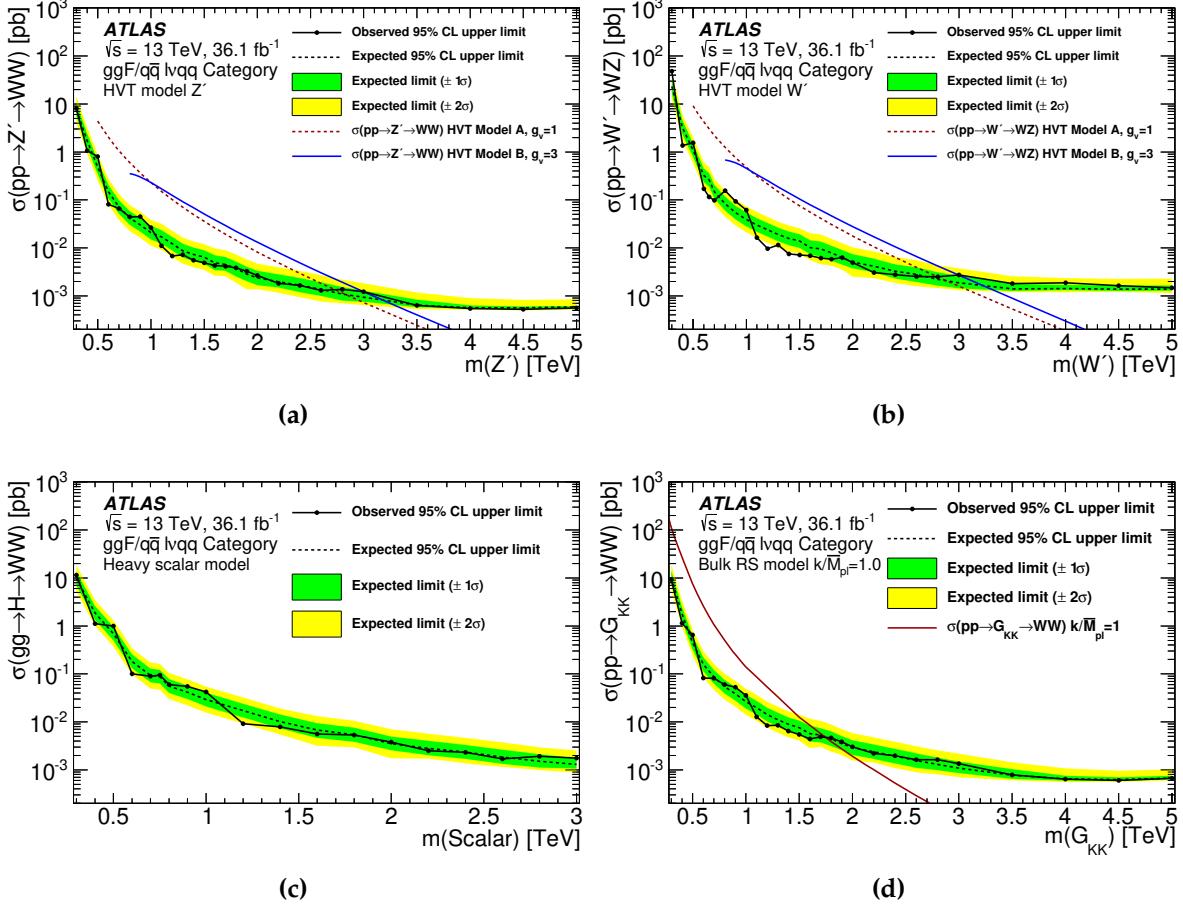
$$\sigma \times BR = \frac{N^{evt}}{\mathcal{L} \times \mu} \quad (4.21)$$

with  $\mathcal{L}$  as the luminosity

The results with the combination of all the signal regions are presented in Fig. 4.4 for the ggF category with theoretical cross-section overlaid together , and Fig. 4.5 for the VBF category. For the W' boson, Z' boson, and the RS graviton, the theoretical cross-section is overlaid together with the expect and observed limits from the experiment.

## 4.4 Combination of VV/VH/ $\ell\ell/\ell\nu$

As mentioned before, the combination of multiple signal regions would help to increase the statistics and improve the measured sensitivity. In addition to the final state this analysis is interested in( $pp \rightarrow WV \rightarrow \ell\nu qq$ ), there are also other analyses which are aiming for the same exotic particles. Therefore, a combination across all the possible final states of those searches was conducted to have a further improvement in the final result. The proposed scheme is to combined the diboson analyses for which the final

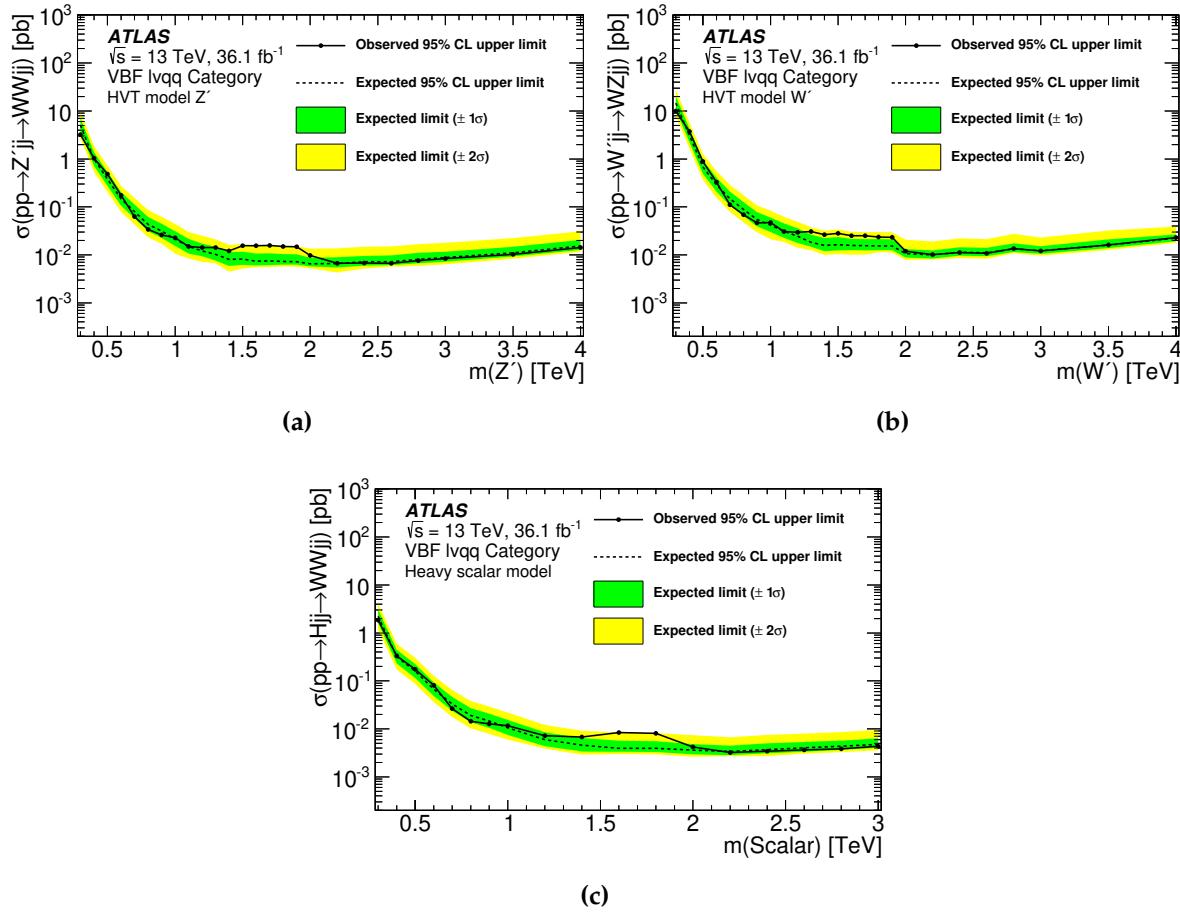


**Figure 4.4:** The limits for the BSM particles via ggF/qqF production. (a) and (b) are for the HVT  $Z'$  and  $W'$  bosons, while (c) is for the NWA scalar boson, and (d) is for the RS graviton.

states of  $VV$  ( $V = W$  or  $Z$  boson) decay are considered to search for the scalar boson, the HVT, and the RS graviton. And, to have a further understanding of the HVT coupling to the SM particles, the dilepton ( $\ell\ell$  and  $\ell\nu$ ) and VH ( $H \rightarrow bb$ ) channels are also taken into the combination.

#### 4.4.1 Combination Strategy

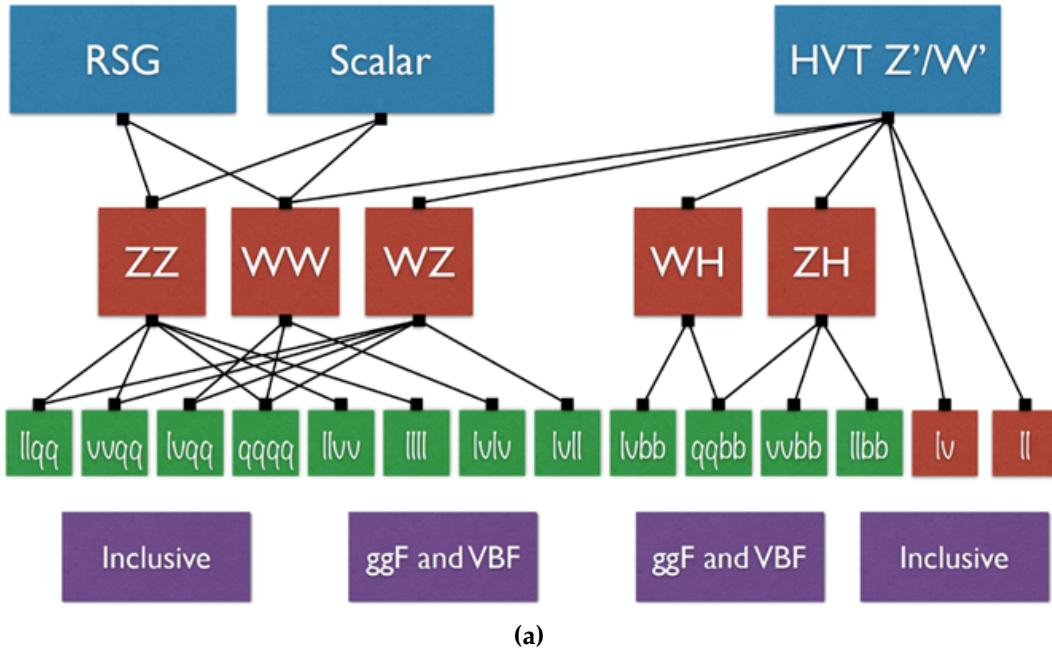
The combination scheme could be seen in Fig. 4.6, and the considered analyses with a brief event selection summary is presented in Tab. 4.2.



**Figure 4.5**

With the number of involved analyses, the likelihood construction of all the final state would be too complicated, so the procedure was conducted step by step. It started from the same medium states of WW, WZ, or ZZ bosons with their fermionic final states, and then, they are integrated into the VV combination. At this stage, the statistic interpretations on RS graviton and NWA scalar boson are completed. Following by that, the VV channels are combined together with VH and dilepton channels to set the limit on mass of  $W'$  and  $Z'$  bosons as well as the coupling strength between the HVT and SM particles.

## Orthogonality



**Figure 4.6:** The scheme for combination of VV, VH, and dilepton analyses with their final states.

Within VV and VH channels, the category orthogonality was kept by the cuts on lepton number,  $E_T^{miss}$ , and b-jet numbers. However, to have the selection on the boson decayed jets, the mass windows were overlapped, so some events went into both VV and VH signal regions. Tab. 4.3 shows the mass windows used in the hadronically decayed bosons. In this case, the events would be given higher priority to go into the VV category and get removed from the VH channels. With the comparison to the original event selection, the expected sensitivity doesn't have significant change which could be seen in Fig. 4.7.

### Nuisance Parameter Correlation

For each individual analysis, more than 100 nuisance parameters are considered. Some of them are commonly applied across the analyses, but there are also the ones which only made the contribution to the dedicated analyses. The following is the list of nuisance parameters which are decorrelated from the other analyses:

**Table 4.2**

Channel	Diboson state	Selection			VBF cat.	Reference
		Leptons	$E_T^{miss}$	Jets		
$qqqq$	$WW/WZ/ZZ$	0	veto	2J	—	[?]
$vvqq$	$WZ/ZZ$	0	yes	1J	—	yes [?]
$\ell vqq$	$WW/WZ$	$1e, 1\mu$	yes	2j, 1J	—	yes [?]
$\ell \ell qq$	$WZ/ZZ$	$2e, 2\mu$	—	2j, 1J	—	yes [?]
$\ell \ell vv$	$ZZ$	$2e, 2\mu$	yes	—	0	yes [?]
$\ell v \ell v$	$WW$	$1e+1\mu$	yes	—	0	yes [?]
$\ell v \ell \ell$	$WZ$	$3e, 2e+1\mu, 1e+2\mu, 3\mu$	yes	—	0	yes [?]
$\ell \ell \ell \ell$	$ZZ$	$4e, 2e+2\mu, 4\mu$	—	—	—	yes [?]
$q q b b$	$WH/ZH$	0	veto	2J	1, 2	— [?]
$v v b b$	$ZH$	0	yes	2j, 1J	1, 2	— [?]
$\ell v b b$	$WH$	$1e, 1\mu$	yes	2j, 1J	1, 2	— [?]
$\ell \ell b b$	$ZH$	$2e, 2\mu$	veto	2j, 1J	1, 2	— [?]
$\ell v$	—	$1e, 1\mu$	yes	—	—	— [?]
$\ell \ell$	—	$2e, 2\mu$	—	—	—	— [?]

**Jet Uncertainties:** The measurement of jets actually have 81 sources of uncertainties, but most of analyses just deploy the simplified schemes for which the 81 sources are combined into 21 or 3 uncertainties. For the analyses using different simplified uncertainty schemes, their jet uncertainties are decorrelated ( $VV \rightarrow \ell v \ell \ell \& \ell \ell v v$ )

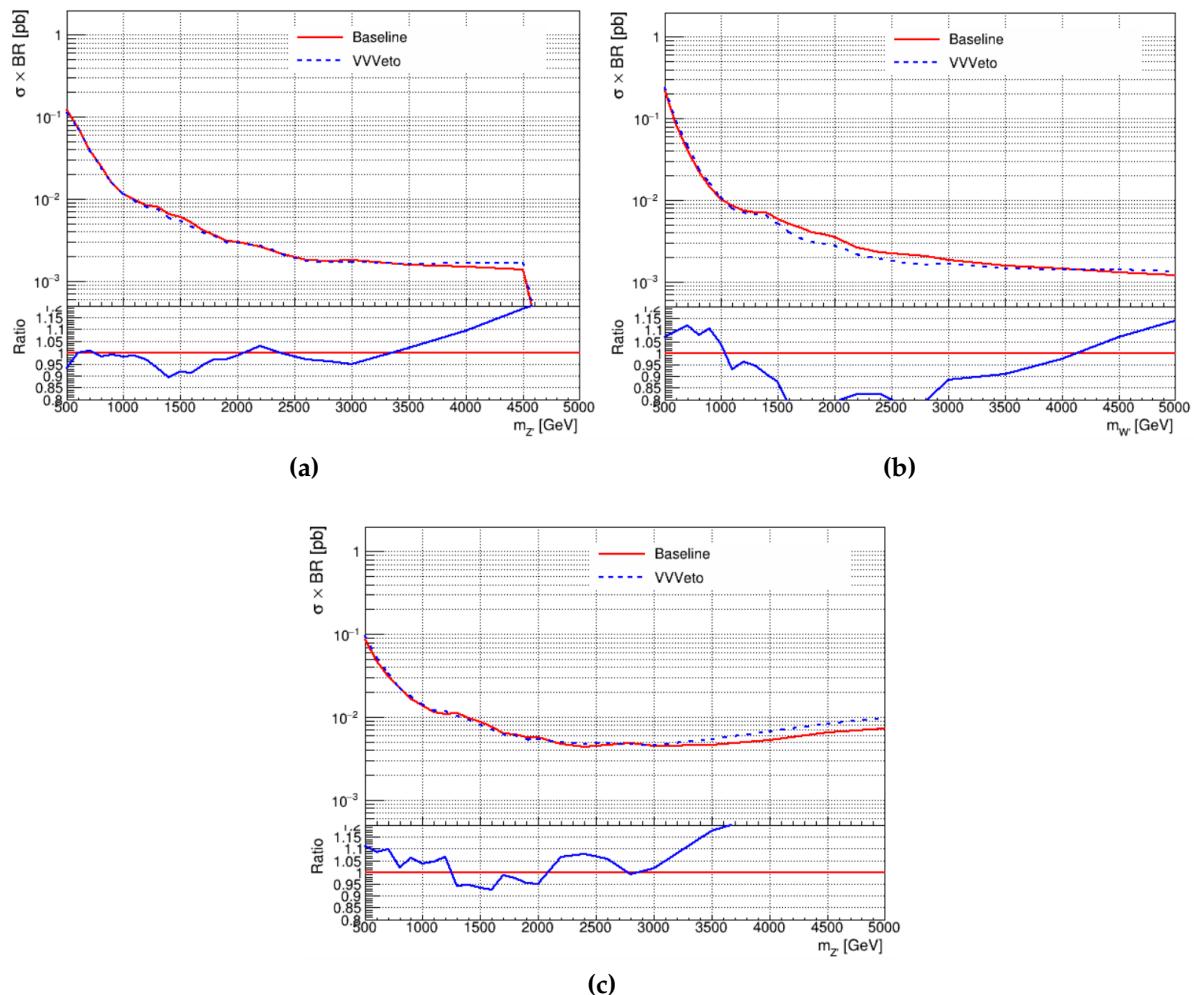
**Electron ID Uncertainty:** The  $VV \rightarrow \ell v \ell v$  analysis has deployed different identification working points in the electron selection, so the related uncertainty is decorrelated.

**Signal and Background Modelling Uncertainties:** The scale factors for the SM background in the likelihood reconstructions are decorrelated as they have different kinematic properties for varied final states. Furthermore, the uncertainties arising from the data-driven estimation are also decorrelated. Furthermore, as the ISR/FSR effect was not considered in the fully leptonically decayed channels, they are decorrelated as well.

**Table 4.3:** The mass windows for the selection on hadronically decayed bosons in VV and VH events

channel	Jet Topo	W	Z	H
$qqqq$	resolved	-	-	-
	merged	[65,95]	[76,106]	-
$\ell\ell qq$	resolved	[62,97]	[70,105]	-
	merged	[65,95]	[76,106]	-
$\ell\nu qq$	resolved	[66,94]	[82,106]	-
	merged	[64,104](LP)	[69,114](LP)	-
$\nu\nu qq$	resolved	-	-	-
	merged	[65,95]	[76,106]	-
$qqbb$	resolved	-	-	-
	merged	-	[70,110] (HP)	[75,145]
$\ell\nu bb/\nu\nu bb$	resolved		[110,140]	
	merged		[75,145]	
$\ell\ell bb$	resolved		[100,145]	
	merged		[75,145]	

#### 4.4.2 Conclusion



**Figure 4.7:** The change in expected limits in the VH channels for (a)  $VH \rightarrow \nu\nu bb$  (b)  $VH \rightarrow \ell\nu bb$  and (c)  $VH \rightarrow \ell\ell bb$



# Chapter 5

## Non-Resonance Searching Strategy

In addition to the physics with resonance particles, the unknown couplings between SM particles is also a portal to new physics. Its signature would be similar to SM, but it might enhance or reduce (interference) the occurrence rate (cross section) for the physical process of interest. However, the deviation from SM prediction would be marginal, so the test on precision measurement could only be achieved with great amount of data.

This analysis is aiming for the phenomenology with “vector boson scattering” (VBS) which has the signature like VBF with one back-to-back high-mass jet pair accompanied by two SM gauge bosons ( $qq \rightarrow VVqq$ ). The phenomenon was predicted by SM and measured in ATLAS Run1 data analysis (2009-2012) with the search for the anomalous quadratic gauge coupling (aQGC), but it doesn’t give a promising result ( $5\sigma$  excess in comparison to null hypothesis) for the existence. This analysis is to extend the search with greater luminosity of data collected in 2015 and 2016 for  $36.1fb^{-1}$ , and the final state of the diboson system is chosen to be semileptonic. This analysis will focus on one lepton channel ( $WV \rightarrow l\nu qq$ ) just like the the resonance search , and the result will be combined with other two semileptonical final states ( $ZV \rightarrow \nu\nu qq/ll qq$ ) for the statistic interpretation.

With the same final state, the object definition was inherited from the resonance search, and the simulation sample and dataset are also reused. However, because the

search is aiming for different signal, the optimization was repeated for the threshold of object and event selection. The most significant change in this analysis is that although  $m_{WV}$  is still reconstructed, it cannot be taken as the discriminant because of no resonance particle in the process. Instead, an algorithm of boosted decision tree (BDT) is performed on the detector variables, and it would give the output of “BDT score” for signal and background separation.

To maximize the sensitivity, the event categorization employs the same strategy to have boosted HP, boosted LP, and resolved regions for signal, W+jet and  $t\bar{t}$  control regions for VBS category only, and the event priority is the same as the resonance search.

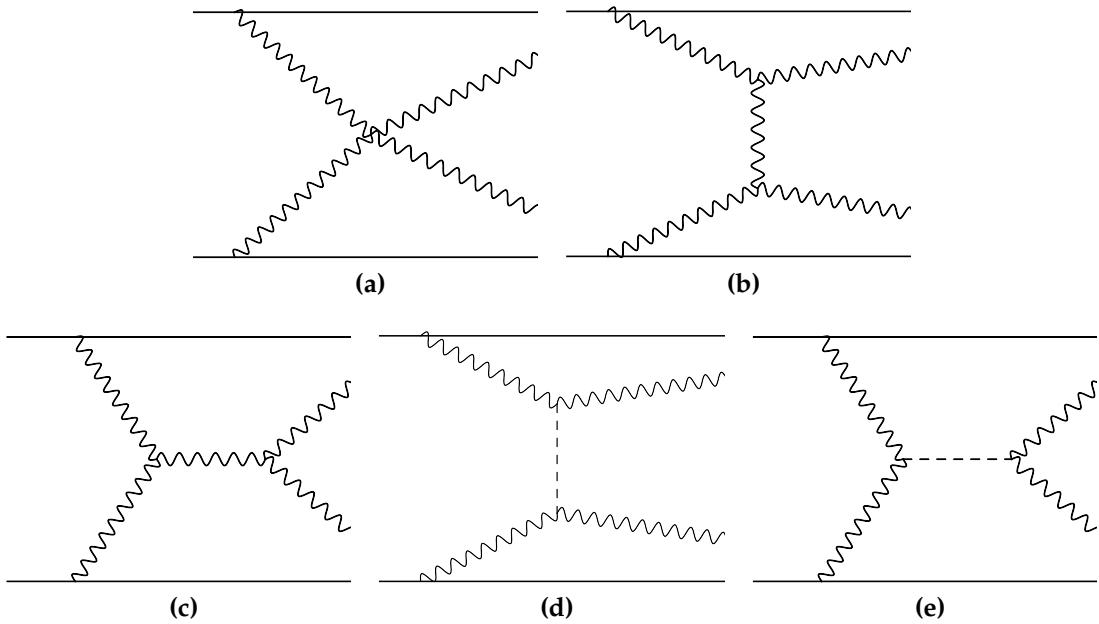
## 5.1 Signal Simulation Samples

Two types of signal signature were generated: SM VBS scattering and anomalous quadratic coupling. As this analysis is a general search for the coupling signature, a couple of physical processes are involved as the signal. In this case, an approximation of effective field theory (EFT) is applied to simplify the simulation.

### 5.1.1 Standard Model Vector Boson Scattering

Under SM, the vector boson scattering is through the coupling to a variety of bosons including W/Z boson, photon, or Higgs boson. The coupling strength is constrained by Higgs mass, so the measurement could be another test on Higgs naturalness and Brout-Englert-Higgs Mechanism. The interactions considered in this analysis are shown in Fig. 5.1 with the order of  $\alpha_{EW}^6$  which also considers the decays of bosons into fermions.

For those interactions, only the longitudinal component of bosons is considered, while the transverse one has relatively low coupling strength, so it is neglected. When the Higgs boson is not involved in the interactions (Fig. 5.1 (a)-(c)), the coupling



**Figure 5.1:** Here are the Feynman diagrams which contribute to the SM VBS signal. The dashed line in figure (b) and (d) are the Higgs boson which couples the interactions. Those interactions are of the order  $\alpha_{EW}^6$  involving the consideration of the decays of the two scattered bosons into fermions.

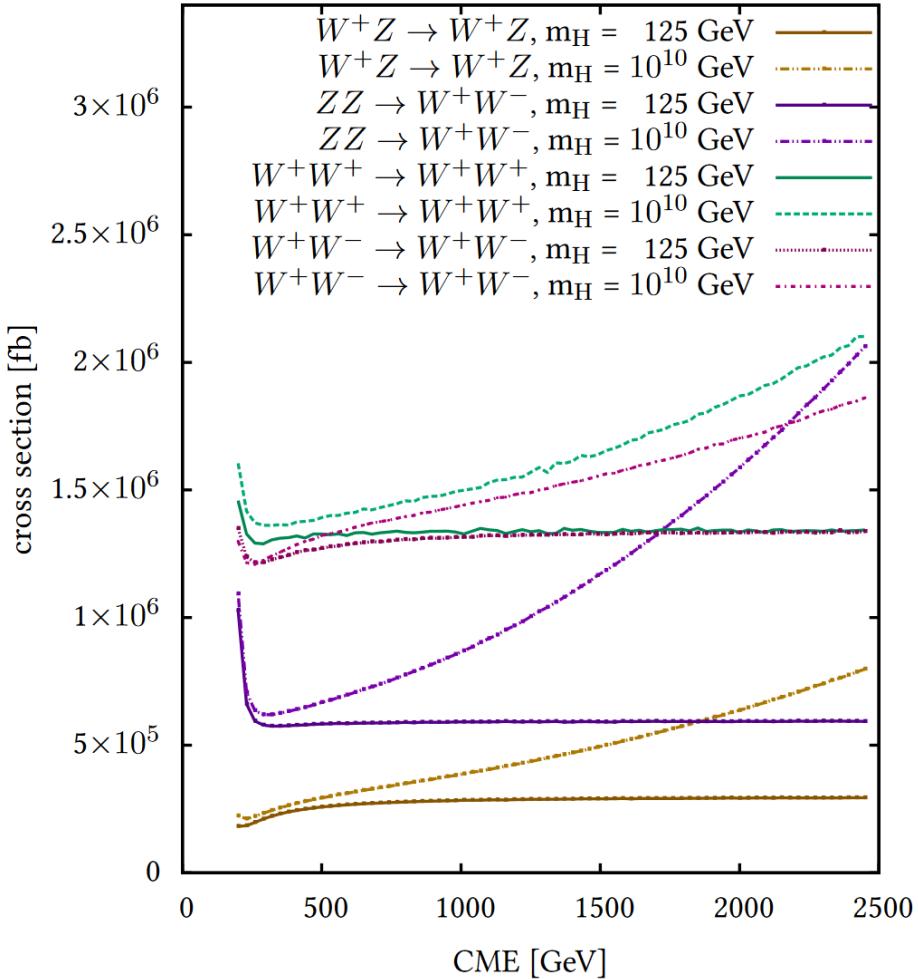
magnitude could be presented in Mandelstam variables as:

$$|\mathcal{M}| = \frac{g^2}{4m_W^2} [s + t] \quad (5.1)$$

with the consideration of only the coupling to W boson for simplicity. This implies that the coupling strength will diverge when energy increases, so the unitarity of  $\rho$  in Eq. 1.23 would be broken due to the enhanced coupling between W and Z bosons. The introduction of the Higgs boson to mediate the boson could put a constraint to prevent the divergence, which makes the coupling magnitude to:

$$|\mathcal{M}| = \frac{g^2 m_H^2}{4m_W^2} \left[ \frac{t}{t - m_H^2} + \frac{s}{s - m_H^2} \right] \quad (5.2)$$

This expression is only valid at the tree level, and the perturbative terms are neglected with the light Higgs boson. However, the high order terms would remain if the Higgs



**Figure 5.2:** The cross-section of vector boson scattering between gauge bosons with Higgs boson of  $125\text{GeV}$  and  $10^{10}\text{GeV}$  as a function of  $\sqrt{s}$ .

mass is above  $1.2\text{TeV}$  with  $\lambda > 4\pi$ :

$$m_H^2 = 2\lambda\nu^2 > 1.2\text{TeV} \quad (5.3)$$

with  $\nu$  as  $246\text{GeV}$  measured from the experiments. This would then make the cross-section diverge again which is shown in Fig. 5.2.

### Sample Production

The signal samples are produced with the setting under SM with Higgs boson mass at  $125\text{GeV}$ . MADGRAPH5\_AMC@NLO v2.3.3 is the chosen generator interfaced by

PYTHIA8 for the fragmentation with the PDF set of NNPDF30LO. The two outcome bosons are required to be on-shell with mass pole from PDG.

As the generation command in the generator is  $pp \rightarrow VVjj$ , so some of the unwanted interaction would also go into the signal samples. Their coupling is still at the order of  $\alpha_{EW}^6$ , but no VBS interaction is involved. With the VBS requirement on event selection, the contribution is well-suppressed.

### 5.1.2 Anomalous Quadratic Coupling (aQGC)

With the light mass of the discovered SM Higgs boson, the Higgs naturalness turns to be an problem. In addition to the BSM heavy Higgs bosons, the hidden couplings between bosons is another approach to this issue. It could lead to the fine-tuning to the SM Lagrange to make high order correction.

To simplify the hidden theory, the approach of EFT is applied which could be presented in Lagrange as:

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \sum_i \frac{C_i}{\Lambda_i^{d-4}} \mathcal{O}_i^d \quad (5.4)$$

where the extended term,  $\sum_i \frac{C_i}{\Lambda_i^{d-4}} \mathcal{O}_i^d$ , is contributed from the anomalous couplings. It is constructed by 3 components:  $\Lambda$  as the energy scale for where the coupling is significant,  $C_i$  as the coefficient of this interaction and  $\mathcal{O}_i^d$  is the operator.  $d$  is used as the number of dimensions of this coupling. With the constraint of  $\Lambda$  in the power of  $d - 4$ <sup>1</sup>, the interactions of higher order could be neglected due to small contribution. When  $\Lambda$  goes to  $\infty$ , that would mean the new physics is unapproachable, and SM would be the only observable phenomenon. This approach has been proven working well to have theoretical agreement to experimental data with the example from Fermi theory.

---

<sup>1</sup> $d - 4$  is applied on the energy scale to keep the dimension consistent in the Lagrangian

The new physics operator,  $\mathcal{O}$ , considered here is based on Eboli model which formulates the new interactions with the components:

- Higgs Field Covariant Derivative:  $D_\mu \Phi = (\partial_\mu + igW_\mu^j \frac{\sigma^j}{2} + ig'B_\mu \frac{1}{2})\Phi$
- Electroweak W Field Covariant Derivative:  $\hat{W}_{\mu\nu} = \sum_i (\partial_\mu W_\nu^i - \partial_\nu W_\mu^i) \frac{\sigma^i}{2}$
- Electroweak B Field Covariant Derivative:  $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$

The notations used here are the same as the ones used in Chapter 1 with  $\sigma^j$  as the Pauli Matrix, and they all have the same dimension of 2:

$$[D_\mu \Phi] = [\hat{W}_{\mu\nu}] = [B_{\mu\nu}] = GeV^2 \quad (5.5)$$

Then, the choice on 4 of them (the same component could be chosen multiple times) are combined into one individual operator with the dimension of 8. 3 type of operators could be categorized by the combinations:

- $D_\mu \Phi$  only: the operators are only composed of  $D_\mu \Phi$  and denoted as  $\mathcal{O}_S^i$  with the free parameters  $f_S^i$ . The index ranges from 0 to 2.
- all the elements: the operators are the mix with all the components denoted as  $\mathcal{O}_M^i$ . The free parameters are denoted as  $f_M^i$  with the index ranging from 0 to 7.
- combination of the electroweak fields: the operators have  $\hat{W}_{\mu\nu}$  and  $B_{\mu\nu}$  which are denoted as  $\mathcal{O}_T^i$  with the free parameters,  $f_T^i$ . The index has the range from 0 to 9.

## Signal Production

it is impossible to investigate all the possible operators, so only one operator of each category is chosen in the signal, which are  $\mathcal{O}_S^0$ ,  $\mathcal{O}_M^0$ , and  $\mathcal{O}_T^0$  with the free parameter,  $f_S^0$ ,  $f_M^0$ ,  $f_T^0$ , while the other operators are tuned to 0. The chosen coupling strength for the free parameters in the simulation is summarized in Tab. 5.1.

Similar to the SM VBS signal, the production is also via MADGRAPH5\_AMC@NLO v2.3.3 interfaced by PYTHIA8 with the PDF set of NNPDF30LO.

**Table 5.1:** Set-up of parameters in Eboli Model for this analysis

	$f_S^0 [10^{-12} TeV]$	$f_M^0 [10^{-12} TeV]$	$f_T^0 [10^{-12} TeV]$
Signal 1	50	0	0
Signal 2	0	5	0
Signal 3	0	0	1

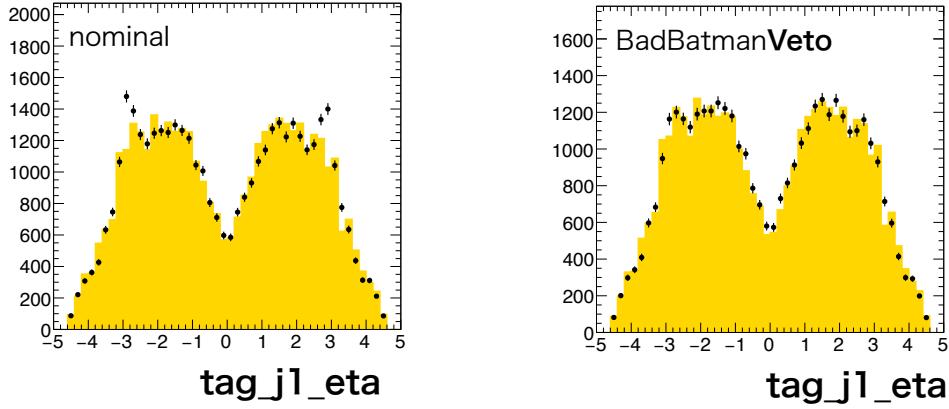
## 5.2 Event Selection

In general, the event selection in this analysis has the same scheme as the resonance search including both trigger and event cleaning. However, as the optimization is conducted with SM VBS signal, some cuts are looser due to the similar kinematic properties with respect to the background. For the same reason, the cuts on topological variables employed in resonance search would also remove significant signal, so they are dropped in this analysis.

Not just the event selection scheme but also the event categorization is adopted here. Three regions are defined with the jet topology as boosted HP, boosted LP, resolved regions (in the order of selection priority) for W+jet control region, top control region, and signal regions (with higher selection priority than the other two). To achieve better sensitivity for aQGC, the dedicated control region is defined with one addition cut in the signal region with  $m_{jj}^{VBS} > 1 TeV$ .

### 5.2.1 Batman Veto

In September 2017, a cell saturation problem in high pile-up runs in 2015 and 2016 was reported in the LAr detector endcap (EMEC). This leads to a large number of low  $p_T$  jets ( $p_T > 20 GeV$ ) falsely reconstructed at  $|\eta| \sim 2.9$  (which is beyond the range where jvt is applicable), and they also made the contribution to  $E_T^{miss}$  reconstruction. Therefore, the events are removed manually by the event and run numbers. Fig. 5.3 shows the jet  $\eta$  distributions before and after the problematic event removal in



**Figure 5.3:** The jet  $\eta$  distribution before (left) and after (right) the problematic event removal

resolved signal region but with 2 loose lepton selected which is another channel along with this 1 lepton analysis.

### 5.2.2 VBS Event Selection

The pair of VBS jets are still chosen to be the one with the highest invariant mass and toward different  $\eta$  direction ( $\eta_1 \times \eta_2 < 0$ ) in a event, but, different to resonance search, the selection shall be conducted after the pair of signal jets from the boson decay (the pair with invariant mass most close to W/Z mass pole) to achieve better sensitivity in resolved region. The full selection of the jet pair is listed below:

- $p_T^{j1} > 30\text{GeV} \&\& p_T^{j2} > 30\text{GeV}$
- $m_{jj}^{VBS} > 400\text{GeV}$
- $\eta_1 \eta_2 < 0$
- no b-tagged

### 5.2.3 Boosted & Resolved Event Selection

Tab. 5.2 and 5.3 are showing the cuts applied to select the events into both signal and control regions. The measurement on both aQGC and SM cross-section doesn't distinguish the processes of WW or WZ, so only one SR is defined for both boosted

**Table 5.2:** Summary of the selection criteria in the definition of the signal region (SR),  $W$ +jets control region ( $W$  CR) and  $t\bar{t}$  control region ( $t\bar{t}$  CR), in the high-purity (HP) and low-purity (LP) categories.

Selection		SR		W CR		$t\bar{t}$ CR	
		HP	LP	HP	LP	HP	LP
$W \rightarrow l\nu$	Num of signal leptons	1					
	Num of vetoed leptons	0					
	$E_T^{miss}$	$> 80 GeV$					
$W/Z \rightarrow J$	Num of large- $R$ jets	$\geq 1$					
	$D_2^{(\beta=1)}$ 50 % WP	pass	fail	pass	fail	pass	fail
	$D_2^{(\beta=1)}$ 80 % WP	—	pass	—	pass	—	pass
	$W/Z$ mass 50 % WP	pass	fail	—	—	pass	fail
	$W/Z$ mass 80 % WP	—	pass	fail	fail	—	pass
Top-quark veto	Num of $b$ -tagged jets	0				$\geq 1$	

and resolved jet topologies.

In comparison to the resonance search, object selections are kept the same like the  $p_T$  threshold or lepton isolation requirements. However,  $E_{miss}^T$  cuts is lowered  $80 GeV$  to enhance the statistics for the training sample into the multivariable analysis. Furthermore, the topological cuts are also removed from this analysis, because they are too stringent for the signal sample. And, the  $b$ -tagging requirement is also changed in the resolved channel with forbidding any  $b$ -tagged jet in the events for signal region.

The definition of  $W$ +jet control regions is still defined by the mass side band of the dijet system in resolved channel and failed mass tagging in boosted channel. With the change on number of  $b$ -tagged jets, the top control region definition is also simplified as the event with any existence of  $b$ -jets.

**Table 5.3:** Summary of the selection criteria of the resolved analysis for the WW and WZ signal regions (SR),  $W+jets$  control region (WR) and  $t\bar{t}$  control region (TR).

cuts		SR	WR	TR
$W \rightarrow \ell\nu$ selection	Number of signal leptons		1	
	Number of veto leptons		0	
	$E_T^{miss}$		$> 80 GeV$	
$W/Z \rightarrow jj$ selection	Number of small jets		$\geq 2$	
	$p_T(j1)$		$> 60 GeV$	
	$p_T(j2)$		$> 45 GeV$	
	$m_{jj}$	[64, 106]GeV	$< 66 GeV$ or [106, 200]GeV	[64, 106]GeV
Top veto	Number of $b$ -tagged jets	0	0	$\geq 1$
Existence of VBF jets		yes		

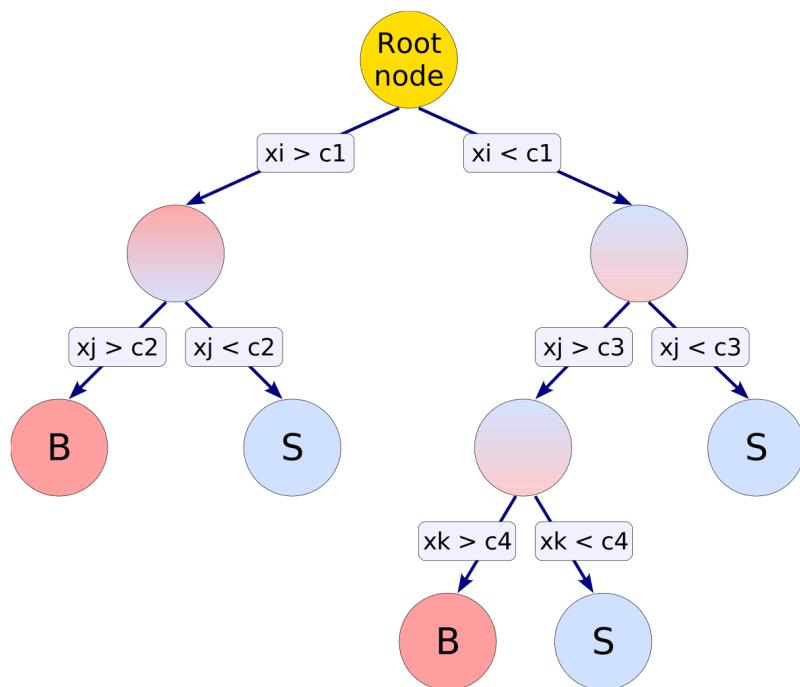
## 5.3 Multivariate Analysis

When the new physics and SM interactions have similar kinematics, their detector signatures would have marginal difference. In this case, no individual variable could be taken as the discriminant to distinguish signal from backgrounds. And, this is why the multivariate analysis is taken into the analysis.

Two candidate algorithms were considered: adaptive and gradient decision tree. However, the outcome of adaptive decision tree is not robust to prevent overtraining. This means the outcome might only have the distinguishing power to the training sample, but the appliance outside is less powerful. Therefore, gradient boost decision (GBDT) is chosen to be used in this analysis.

### Decision Tree

Decision tree is constructed with a series of binary decisions. Those decisions are made by whether the events could pass “a” cut which can give the best prediction. Two types of decision tree are defined by signal and background separation. The full scheme could be presented as Fig. 5.4. To make the best decision on signal and



**Figure 5.4:** The scheme of a decision tree. The nodes are where the decision are made by whether the event could pass the cut on  $x_i$ ,  $x_j$ , and  $x_k$ . The bottom nodes are the final outcomes labeled as S for more signal events categorized into the node, and B for more background events

background separation, the *Gini Index* is defined as:

$$I_G = p(1 - p) \quad (5.6)$$

where  $p$  is the percentage of signal events (purity) in the node. The decision would then be made to optimize the increase of  $I_G$  in the mother node:

$$Gain = I_G^{\text{mother}} - \frac{n_1}{N} I_G^{\text{daughter1}} - \frac{n_2}{N} I_G^{\text{daughter2}} \quad (5.7)$$

where the two daughter nodes are split from the mother node (event number,  $N$ , divided to  $n_1$  and  $n_2$ ), and their  $I_G$ 's are reweighted. After repeating this procedure, the events get into the leaf nodes where they are labeled by signal or background depending on which kind of events takes over bigger proportion. To avoid over-training, a further procedure called “pruning” is applied. This procedure is conducted from the bottom, and the node with little increase of the separation power (*Gini Index*) is removed. If the following branches are all removed, the mother node making the decision would be turned into another leaf node.

The other decision tree is also used in this analysis which is called “regression tree” for the prediction of a truth value which is not in the binary format. In this case, the *Gini Index* is not available, and, instead, the average squared error is used:

$$\sigma = \frac{1}{N} \sum_i^N (\hat{y}_i - y_i^{truth})^2 \quad (5.8)$$

with  $\hat{y}_i$  as the average truth value of the events in the same node and  $y_i^{truth}$  as the truth value of the events itself. The decision is made with a cut which makes  $\sigma$  as the minimum.

In general, Decision tree has the advantage of being handy for use, easy to understand, and flexible, but it also has the disadvantage of being weak to statistic fluctuation and correlation between input variables especially when only one tree is trained. Therefore, there are a couple of methods to train trees into “forests” like random forest, and the one used in this analysis is “Gradient Boost”.

## Gradient Boost

Due to the insufficiency of decision tree, the approach called “boost” is applied into the sample events after each tree is grown up. The process is determined by deviation of the predicted value from the truth value with a “Loss Function” defined as:

$$L(F(x_i), y_i) = \sum_i^n \ln(1 + \exp(-2F(x_i)y_i)) \quad (5.9)$$

where  $F(x_i)$  is the predicted value from the outcome of decision trees for the  $i$ th event, and  $y_i$  is its truth value. The training process is then conducted in the following steps:

- (a) The first tree is built up with the procedure as mentioned above to optimize the *Gini Index*.
- (b) The predicted value of each event,  $F_0(x_i)$ , is then assigned by the type of nodes.  
(signal node:  $F_0(x_i) = 1$ , background node:  $F_0(x_i) = -1$ )
- (c) The gradient of Loss Function,  $\partial L(F_0(x_i), y_i) / \partial F_0(x_i)$ , is calculated for each event.
- (d) The average of the gradient from all the events in a node is assigned as a new label.
- (e) Instead of using the label of background (-1) and signal (+1), the average of gradient is taken as the new label on the events for next training
- (f) The training is conducted as a regression tree by minimizing the average squared error when making each decision
- (g) After the tree is constructed, the average expected value of each leaf node,  $\hat{y}_i$ , is taken to correct  $F_0(x_i)$ :

$$F_1(x_i) = F_0(x_i) + \beta \hat{y}_i \quad (5.10)$$

with  $\beta$  as the training rate. Smaller  $\beta$  gives better precision, but the training also needs more trees to achieve the required accuracy.

**Table 5.4:** The customized parameters in the Gradient Boosted Decision Tree Configuration

Option	Defined Value	Description
NTrees	800	Number of trees
MinNodeSize	5%	Minimum percentage of training events required in a leaf node, or the leaf shall be trimmed off
Shrinkage	0.3	Training rate used in this analysis
nCuts	20	Number of grid points in variable range used in finding optimal cut in node splitting
MaxDepth	4	Max depth of the decision tree allowed
BaggedSampleFraction	0.5	The fraction of events in the full sample for training of each tree.

(h) step (c) to step (g) are then repeated until  $N$  tree are built.

The final outcome,  $F_N(x_i)$ , of this training is then called “BDT score” which is used as the final discriminant for this analysis.

## Training

Tab. 5.4 is showing the customized parameter configuration in the training for this analysis, and the ones not mentioned are taken as default value. It could be noted that not all events in the training are used in each tree, but a random sampling (“bagging”) is employed. This is to smear the statistic fluctuation to avoid over training.

In this analysis, the SM VBS samples is taken as signal, and the background sample is composed of simulated  $t\bar{t}$  and W+jets interactions. Both of them are split into two samples with equal size: the events of even event numbers are taken as sample A, and odd event number ones are taken as sample B. Then, two trainings are conducted on both samples and evaluated on the other one. Afterwards, data are processed with both of the two trainings also with the event splitting by event numbers. The final result was given as the combination of the two trainings. With drastically different kinematics, the training is performed on merged and resolved channels respectively.

At the very beginning, the trainings were conducted with more than 50 input variables including 4-vector of the selected objects, topological parameters between them, and various associated observables like jet-width. Only the subsets of uncorrelated variables showing great separation power were chosen. Two variables were newly defined here:

- boson centrality ( $\zeta_V$ ): This is to identify the topology that the two scattered bosons are supposed to be within the  $\eta$  gap of the two scattered jets:

$$\Delta\eta_- = \min(\eta(V_{had}), \eta(V_{lep})) - \min(\eta(j_1^{VBS}), \eta(j_2^{VBS})) \quad (5.11)$$

$$\Delta\eta_+ = \max(\eta(V_{had}), \eta(V_{lep})) - \max(\eta(j_1^{VBS}), \eta(j_2^{VBS})) \quad (5.12)$$

$$\zeta_V = \min(\Delta\eta_-, \Delta\eta_+) \quad (5.13)$$

- jet width: the jet calorimeter width defined as the  $p_T$  averaged distance of calorimeter clusters to the jet axis:

$$width = \frac{\sum_i \Delta R(j, c^i) p_T(c^i)}{\sum_i p_T(c_i)} \quad (5.14)$$

with  $c^i$  representing the cluster entities inside the jet.

Tab 5.5 is presenting the variable importance for BDT training with definition as the percentage of variables used to make a decision.

## Result

The final result with the output of BDT response ( $F_{x_i}$ ) is shown in Fig. 5.5. As mentioned above, signal events would have the response close to 1, while background ones tend to have the outcome of -1.

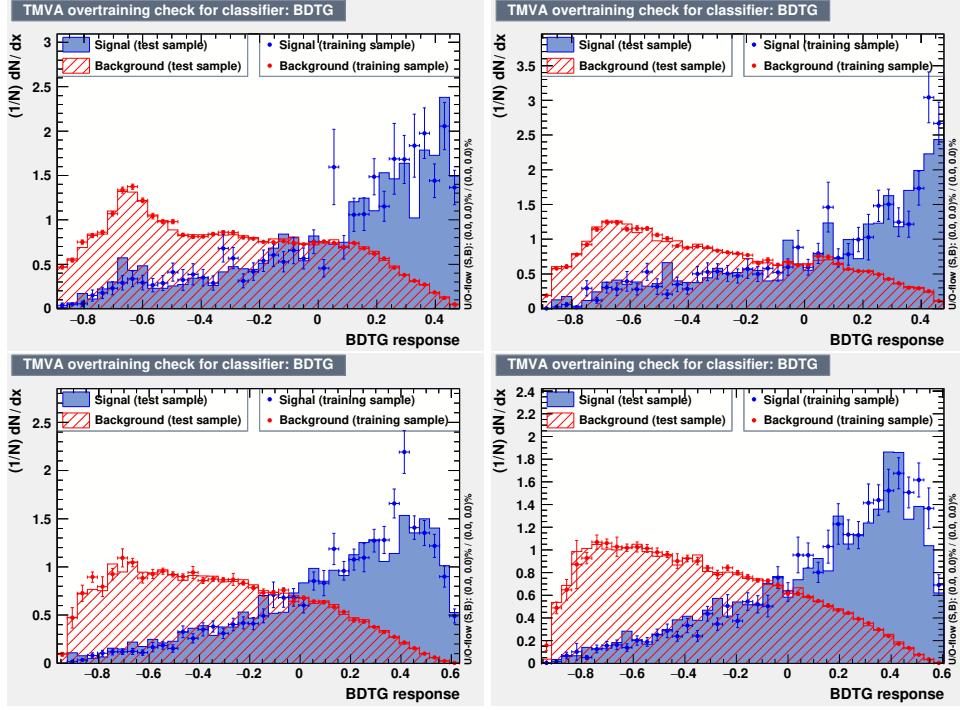
**Table 5.5:** The input variable importance in Gradient BDT training

SelectionVariable	Importance	
	Resolved	Merged
$M_{VV}{}^{tag}{}^{tag}$	$1.532 \times 10^{-1}$	$4.726 \times 10^{-1}$
$p_T^{sig\ j_2}$	$8.833 \times 10^{-2}$	NA
$\eta(\ell)$	$7.939 \times 10^{-2}$	$1.366 \times 10^{-1}$
$\zeta_V$	$7.866 \times 10^{-2}$	$1.970 \times 10^{-1}$
$width(sig\ jet_2)$	$7.151 \times 10^{-2}$	NA
$width(tag\ jet_1)$	$6.933 \times 10^{-2}$	NA
$width(sig\ jet_1)$	$6.354 \times 10^{-2}$	NA
$p_T^{tag\ j_1}$	$6.166 \times 10^{-2}$	NA
$\Delta\eta(j_1, j_2)$	$6.017 \times 10^{-2}$	NA
$\Delta R(\ell, \nu)$	$5.450 \times 10^{-2}$	NA
$p_T^{tag\ j_2}$	$5.352 \times 10^{-2}$	$1.939 \times 10^{-1}$
$N_{trk}(sig\ jet_1)$	$5.102 \times 10^{-2}$	NA
$width(tag\ jet_2)$	$4.093 \times 10^{-2}$	NA
$N_{trk}(tag\ jet_1)$	$3.607 \times 10^{-2}$	NA
$N_{trk}(sig\ jet_2)$	$2.121 \times 10^{-2}$	NA
$N_{trk}(tag\ jet_2)$	$1.697 \times 10^{-2}$	NA

## 5.4 Background Modeling

The modelling strategy is similar to the resonance search. The two dominant background interactions,  $t\bar{t}$  and W+jets, are constrained using dedicated control regions, while the other subtle contributions are without constraint in the signal region fitting. However, to achieve higher precision measurement, some of the cuts are loosened. In this case, if the mismodelling of  $m_{jj}^{VBS}$  is taken into the normalization fitting, it leads to huge uncertainty and degrade the sensitivity to measurement. Therefore, an extra event reweighting is applied on the W+jet MC events in the control region.

The multijet background contribution in the non-resonance search is higher than the resonance analysis due to the lack of anti-QCD cuts from the topological variables. It is also estimated with the same fake factor from resonance search, because the distri-

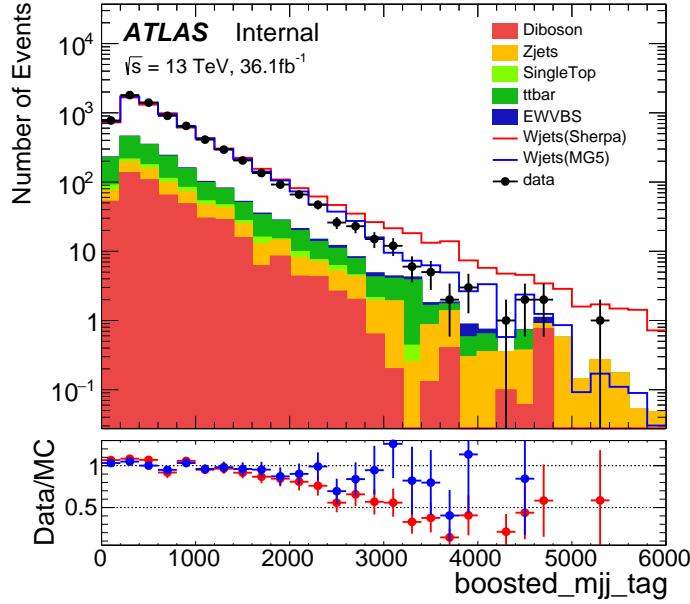


**Figure 5.5:** Comparison of test and training BDTG response distributions in 1-lepton channel, for merged (top) and resolved (bottom) regimes. Results obtained with even (odd) event numbers used for training are shown on left (right).

bution shape is supposed to remain the same with similar final state. The comparison between data and background (pre-fit) will be presented in next chapter along with the post-fit distribution.

### 5.4.1 VBS $m_{jj}$ Modelling

As what was observed in the resonance search, an unknown issue is underlying in the simulation for W+jet events with Sherpa. With the comparison to data, a slope could be seen on the ratio of data to simulation in  $m_{jj}^{VBS}$  distribution shown in Fig. 5.6, and Madgraph sample gives better agreement. However, Sherpa sample has more events for statistics, so it is chosen for the estimation on W+jets background. To remodel the distribution, an extra weight is derived with  $m_{jj(J)}^{VBF}$  distribution in W+jets control



**Figure 5.6:** Comparison of data and W+jets MC samples from Sherpa and Madgraph 5 in the boosted W+jets control region.

region:

$$w(m_{jj}^{VBS}) = \frac{N^{data} - (N^{mc} - N^{mc(W+jets)})}{N^{mc(W+jets)}} \quad (5.15)$$

The estimation is performed as a linear fitting in  $m_{jj(J)}^V$  bins:

$$m_{jj(J)}^V = [50, 60, 70, 100, 150, 200, 300] [GeV] \quad (5.16)$$

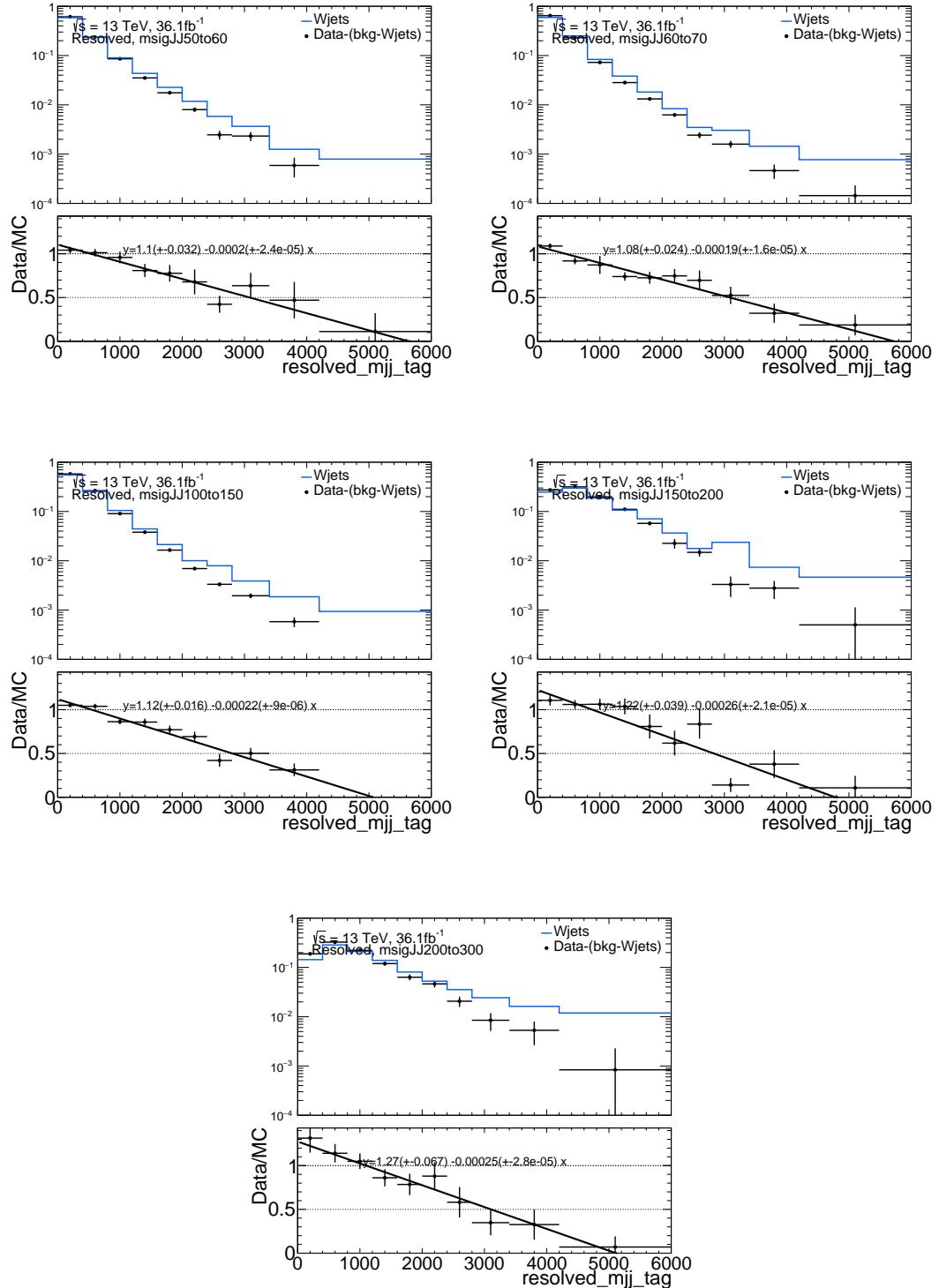
where the bin of 70-100 GeV is removed because it is defined as signal region. The fitting in multiple bins is to investigate the weight dependence on  $m_{jj(J)}^V$ . The result of fitting could be seen in Fig. 5.7 and 5.8. With consistent result between each  $m_{jj(J)}^V$  slice, it was determined to single fitting function for the reweighting, and the parameters employed is shown in Tab 5.6 with  $1\sigma$  uncertainty from statistical fluctuation. As the

discrepancy was only seen in  $m_{jj}^{VBS}$ , the validation was performed by applying the weight in the other distributions, and the result is in Fig. 5.9 which is showing the

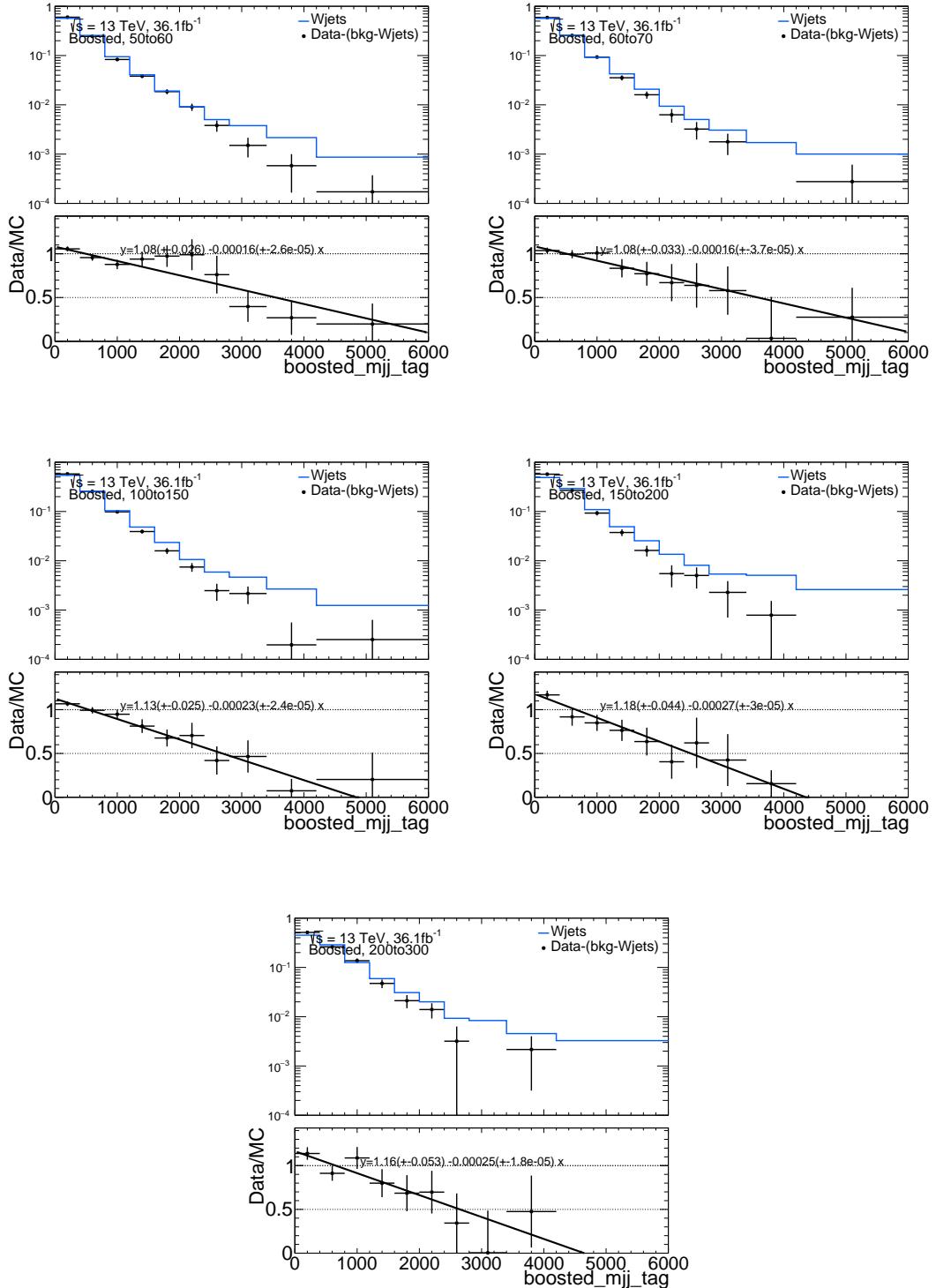
**Table 5.6:** Estimated  $m_{jj}^{VBS}$  reweighting functions for W+jets events.

Fitting parameters	Resolved	Merged
$p_0$ (constant)	$1.1 \pm 0.04$	$1.1 \pm 0.02$
$p_1$ (slope) [ $GeV^{-1}$ ]	$-0.00021 \pm 0.00002$	$-0.00019 \pm 0.00003$

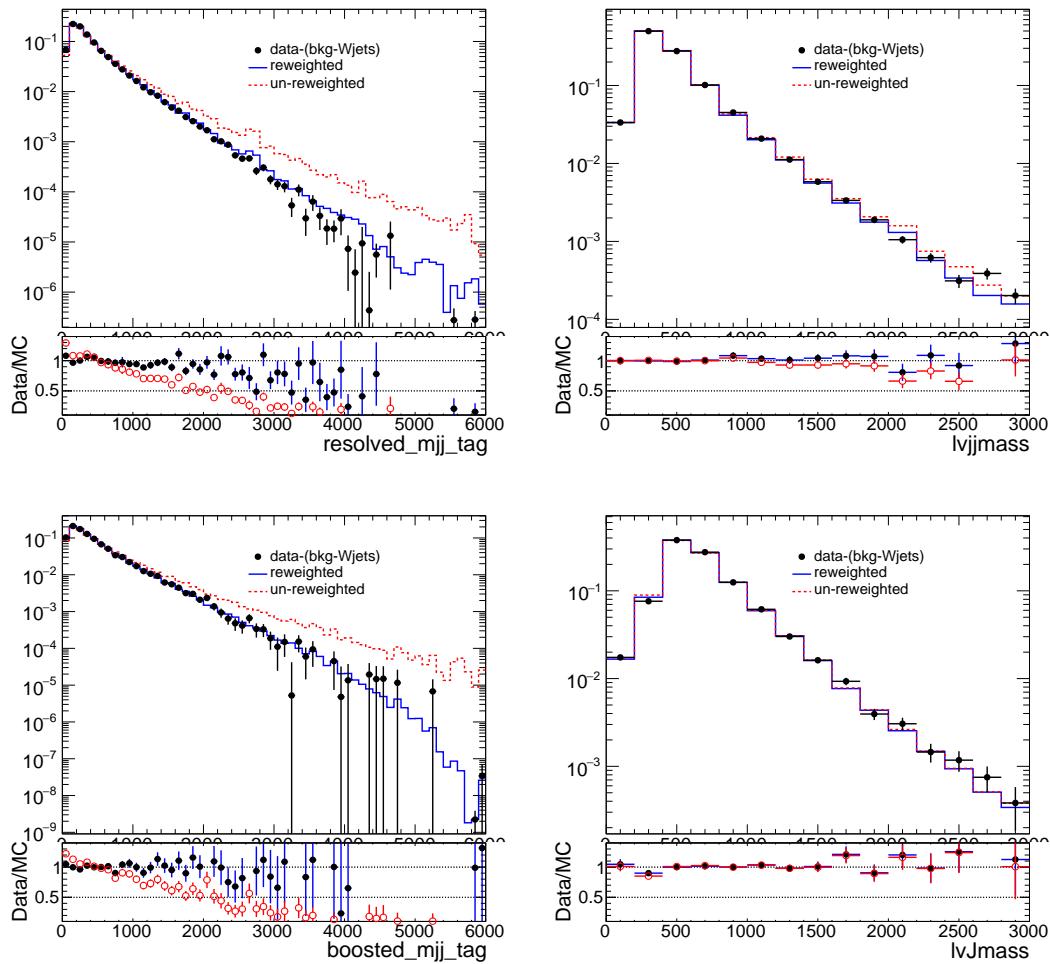
agreement is significantly improved for  $m_{jj}^{VBS}$  distribution with little disturbance on the other kinematic properties.



**Figure 5.7:** Fit of  $m_{jj}^{VBS}$  slope in W+jets resolved CRs, in different slices of  $m_{jj}^V$ .



**Figure 5.8:** Fit of  $m_{jj}^{VBS}$  slope in  $W+jets$  boosted CRs, in different slices of  $m_J^V$ .



**Figure 5.9:** Comparison of  $m_{jj}^{VBS}$  and  $m_{VV}$  distributions before and after the  $m_{jj}^{VBS}$  reweighting for events in the resolved (top) and boosted (bottom) W+jets control regions