Diboson resonance searches in the final state of the $\ell\nu qq$ with improved large-R jet uncertainties in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

Author:
Takuto KUNIGO

Abstract

A number of theories proposed for extensions of the Standard Model (SM), such as composite Higgs models, warped extra dimensions, models with an extended Higgs sector, and grand unified theories, predict diboson resonances. While the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) presented searches for diboson resonances in various decay channels, no evidence of such resonances has been observed.

This thesis presents a search for a resonance, in a mass range from 300 GeV to 5 TeV that decays into a WZ or WW boson pair. Two different production modes are considered in this thesis: (1) the vector-boson fusion (VBF) and (2) the gluon-gluon fusion (ggF) or quark-antiquark ($q\bar{q}$) annihilation. We consider the semileptonic final state where one W boson decays leptonically ($W \to \ell\nu$ with $\ell = e, \mu$) and the other W/Z boson (denoted by V) decays hadronically ($V \to q\bar{q}'/q\bar{q}$ with q, q' quarks). For the high-mass resonances, opening angles between the quarks from the hadronically decaying V bosons are small; the two quarks can be identified as a single jet with a large radius in this case. In contrast, we reconstruct the hadronically decaying V bosons as two separate jets for the low-mass resonances. The major sources of the SM background are (i) top quark pair production and (ii) W boson production associated with hadron jets. These background effects are normalised by the fit results in a set of dedicated control regions.

No significant excess is found in this search. Limits on production cross-section times the branching ratio are obtained as the spectrum of the reconstructed invariant mass of the WV resonance candidates, $m(WV)$, arising from three signal models: an additional heavy Higgs boson predicted by many theories beyond the SM, a heavy vector triplet parameterization based on a simplified phenomenological Lagrangian, and a bulk Randall-Sundrum model. The data exclude a heavy vector triplet with mass below 2750 ˙GeV in a weakly coupled scenario and 3000 GeV in a strongly coupled scenario, as well as a Kaluza-Klein graviton with mass below 1750 GeV.

In order to improve sensitivity to the diboson resonances, we develop a set of data-driven calibrations to the hadronically decaying V bosons that are highly boosted. By these calibrations, we reduce the systematic uncertainties on the jet energy scale by approximately 2%. The expected improvements in the search for diboson resonances by applying the new calibrations are also estimated.

Acknowledgements

I would like to express my deepest gratitude to Prof. Tsuyoshi Nakaya and Prof. Masaya Ishino, whose comments and suggestions were of inestimable value for my Ph.D studies. I also owe a very important debt to my advisers Dr. Toshi Sumida and Dr. Takuya Nobe. I could not have imagined having any better advisers and mentors for my Ph.D studies.

I would express my appreciation for all my colleagues in the Diboson resonance search with the $VH \to \ell\nu qq$ final state group, in particular Dr. Chunhui Chen and Dr. Kalliopi Iordanidou, who are the contact editors for the analysis, and Mr. Ismet Siral, Ms. Natasha Lee Woods, and Mr. Robert Les, co-analysers of my diboson resonance search analysis. I would acknowledge to the *in-situ* Large-*R* jet calibration team, especially to Dr. Mario Campanelli, Dr. Matthew Edgar Leblanc, Dr. Marcel Vos, Dr. Steven Schramm, Dr. Bogdan Malaescu, Dr. Pierre-Antoine Delsart, Dr. Jeffrey Rogers Dandoy, Mr. Matthew Berg Epland, Ms. Eva Hansen, Mr. Tatsumi Nitta. I would like to offer my special thanks to the ATLAS TGC group and the ATLAS Japan group, especially to Prof. Osamu Sasaki, Prof. Hisaya Kurashige, Prof. Kohei Yorita, Dr. Masato Aoki, Dr. Yasuyuki Okumura, Dr. Koji Terashi, Dr. Yuji Enari, Dr. Tatsuya Masubuchi, Dr. Tomoyuki Saito, Dr. Tomoe Kishimoto, Dr. Makoto Tomoto, Dr. Yasuyuki Horii, Dr. Junpei Maeda and Dr. Kentaro Kawade.

I would like to offer my thanks to CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. I wish to extend my thanks to all collaborators of the ATLAS experiment. I would like to thank JSPS KAKENHI Grant Number 15J02931 for a grant that made it possible to complete this study. I would express my gratitude to Prof. David Francis for accepting me during my stats at CERN in the JSPS program. This work was also supported financially by High Energy Accelerator Research Organization (KEK) and the ATLAS Japan groups. I would acknowledge to Prof. Shoji Asai and Prof. Kazunori Hanagaki, who are the representatives of the groups.

I want to express my cordial appreciation also to the other members of our laboratory: Dr. Atsuko Ichikawa, Dr. Osamu Tajima, Dr. Roger Wendel, Dr. Tatsuya Kikawa, Dr. Shunsuke Adachi, Dr Shuhei Obara, Dr. Shunsuke Honda, Mr Shigeto Seki, Ms. Ina Sanjana, Mr. Takuya Tashiro, Mr. Ichinori Kamiji, Mr. Kota Nakagiri, Mr. Keigo Nakamura, Mr. Jiang Miao, Mr. Satoshi Shinohara, Mr. Sei Ban, Mr. Shunichi Akatsuka, Mr. Yosuke Ashida, Mr. Shunsuke Tanaka, Mr. Yohei Noguchi, Ms. Ayami Hiramoto, Mr. Kesuke Kondo, Mr. Koichiro Kuniyoshi, Mr. Junta Komine, Mr. Kenji Yasutome, Mr. Tomohumi Abe, Mr. Takuji Ikemitsu, Mr. Takahiro Odagawa, Mr. Soichiro Kuribayashi, Mr. Masanori Tajima, Mr. Masayuki Hatano, and Mr. Yuya Mino.

Finally, I would also like to express my warmest gratitude to my family in particular my mother and my deceased father, Tomoko and Yoshihiro, for their moral support and warm encouragements. You have inspired me to pursue my dreams, and afforded me every opportunity to succeed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Particle Physics has progressively established the Standard Model (SM), which describes the fundamental constituents of matter and their interactions governed by three of the four fundamental forces. With the discovery of the Higgs boson announced by the ATLAS and CMS collaborations at LHC [1, 2], the SM was confirmed again as a successful description of the nature. The SM can give a convincing explanation up to the electroweak scale, however, we already know some problems existing that cannot be explained by the SM, for example, the hierarchy problem [3], dark matter, dark energy, and the gravity.

1.1 Unsolved problems in the standard model of particles physics

The hierarchy problem is a question which asks why the Higgs boson mass (125 GeV) is much smaller than the Planck mass (1.2×10^{19} GeV). The Higgs boson receives quadratic corrections from all the particles, thus its mass should be tremendously large unless there is an incredible fine-tuned cancellation between the quadratic radiative corrections and the bare mass. Although the fine-tuning is not a fundamental problem of the SM, it seems rather unnatural that the cancellation of the loop corrections occur with a precision of the order of 10^{-17} [3].

Unlike normal matter, the dark matter does not interact with observable electromagnetic radiation, thus can not directly be detected. Whereas the existence of dark matter is well-established by many astronomical measurements [4], its nature still remains one of the unsolved problems of particle physics. The dark energy makes up approximately 68 % of the universe in terms of its energy(mass) [4]. It is confirmed by cosmological measurements, for example cosmic microwave background [5]. We have not understood about the dark energy except for it makes accelerated expansion of the universe.

The gravity is not taken into account in the SM, but we know its existence. If we attempt to describe the gravitational interaction by a SM-like Lagrangian in the quantum field theory, the intermediating particle, graviton, is expected to be massless, and must have spin-2.

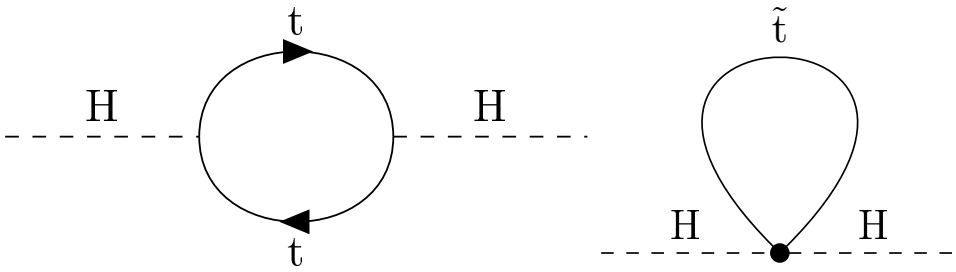
1.2 New physics in diboson resonances

To solve the problems in the SM listed above, enormous number of new models beyond the SM (BSM) have been proposed. For example, Super-Symmetry (SUSY) is one of the most sophisticated extension of the SM, which introduces a new symmetry and predicts a partner for each particle in the SM. With these SUSY particles, it is possible to solve the hierarchy problem; the contributions of the SM particles to the Higgs mass (Figure 1.1(a)) are cancelled by the contributions of the SUSY particles (Figure 1.1(b)), [6–8].

This thesis, however, will not focus on specific theories but a distinct signature which is expected in various BSM theories. One of the most prominent signature is diboson production (WW,WZ and ZZ) resulting from the decay of a new particle predicted in the BSM theories, such as composite Higgs models [9, 10], warped extra dimensions [11–13], models with an extended Higgs sector [14, 15], and grand unified theories [16–18]. In this section, I will explain the physics motivation of the diboson resonance searches, then overview the candidate BSM theories.

1.2.1 Physics motivation of diboson resonance searches

In the absence of a scalar boson, the scattering amplitude of longitudinally polarised vector bosons (Vector Boson Scattering, VBS) would violate unitarity at centre-of-mass energy of order 1 TeV [19, 20]. The unitarity for VBS



(a) (b)

FIGURE 1.1: One-loop quantum correction to the Higgs squared mass. (a) The most significant correction to the Higgs mass originates from the heaviest particle, top quark. (b) These corrections can be cancelled by a superpartner particle, $\tilde{t}$ for top quark.

Process	SM prediction	Observed value	Dataset, Reference
$pp \to W^{\pm}W^{\pm}jj \to \ell\nu\ell\nu jj$	$3.08^{+0.45}_{-0.46}$ fb	2.91 ± 0.51(stat) ± 0.27(syst) fb	$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}, ATLAS [22]
$pp \to W^{\pm}W^{\pm}jj \to \ell\nu\ell\nu jj$	4.25 ± 0.27 fb	3.83 ± 0.66(stat) ± 0.35(syst) fb	$\sqrt{s}$ = 13 TeV, 35.9 fb^{-1}, CMS [23]
$pp \to W^{\pm}Zjj \to \ell\nu\ell\ell jj$	0.13 ± 0.01 fb	(0.63 fb, upper limits at 95 % confidence level)	$\sqrt{s}$ = 8 TeV, 20.3 fb^{-1}, ATLAS [24]
$pp \to ZZjj \to \ell\ell\ell'\ell' jj$	$0.29^{+0.02}_{-0.03}$ fb	$0.40^{+0.21}_{-0.16}(\text{stat})^{+0.13}_{-0.09}(\text{syst})$ fb	$\sqrt{s}$ = 13 TeV, 35.9 fb^{-1}, CMS [25]

TABLE 1.1: Vector boson scattering cross-section summary. The ATLAS and CMS collaborations performed searches for VBS notably in the same-sign WW channel. The observed cross-section values agree with the SM prediction values, whereas the statistical uncertainties are currently large.

is restored by exchanging the Higgs boson in the SM framework. After the discovery of the Higgs boson at LHC [1, 2], the most urgent task for experimental physicists is to test whether the Higgs boson is the only responsible for unitarity restoration for VBS. Any deviation from the SM breaks the cancellation between the VBS amplitudes (Figures 1.2(a)-1.2(c)) and amplitudes that involve the Higgs boson (Figures 1.2(d), 1.2(e)). Thus the final test of the nature of the Higgs boson can be carried out by precision measurements of the scattering of two vector bosons, $VV \to VV$ with $V = W$ or Z, permitting the final test of the electroweak symmetry breaking (EWSB).

Both the ATLAS and CMS collaborations performed searches for VBS using LHC data in the same-sign WW channel, WZ channel, and ZZ channel as summarised in Table 1.1. The difference of the measured cross-section between the ATLAS and CMS results in the same-sign WW channel arises from the difference in the requirements on the objects, which are summarised in Table 1.2. In ATLAS, we apply slightly tighter requirements both on the leptons and jets compared to CMS. All the results on the VBS measurements are dominated by statistical error even using the LHC full data. Furthermore, the current VBS searches exploit only the final state where both vector bosons decay leptonically, called 'fully leptonic' final state. The semi-leptonic final state, in which one vector boson decays leptonically and another boson decays hadronically, has a significantly larger branching ratio, however, this channel is difficult to be studied mainly due to two reasons; (1) the difficulty of reconstruction of hadronically decaying W and Z bosons, and (2) irreducible background events (W/Z+jets and $t\bar{t}$)[1]. Thus we started vector boson resonance searches as a first step, aiming to develop a key technique; large-R jet energy calibration to improve reconstruction efficiency of the hadronically decaying W and Z bosons.

The final goal of this study is to test the EWSB in the VBS processes, however, the vector boson resonance search presented in this thesis is also critical to be studied because many BSM models, summarised in this section, predict new particles decaying into a pair of vector bosons.

[1]The CMS collaboration adopts a different method from us. They use a 2D maximum likelihood fit [21]. This method may be able to provide better background estimation.

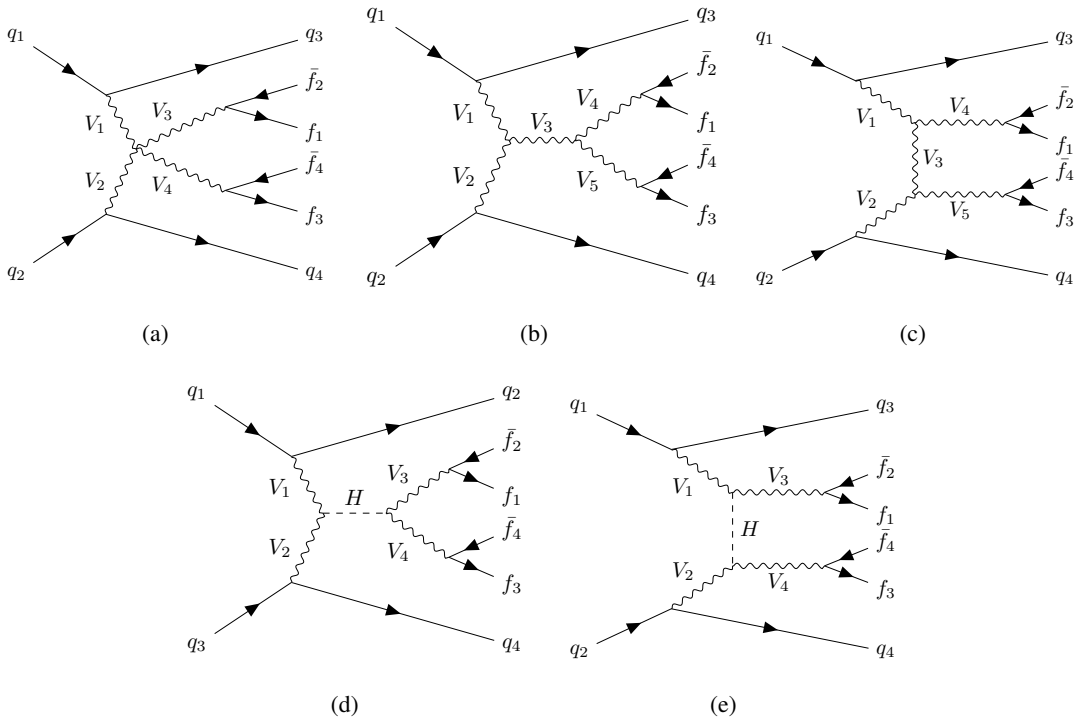


FIGURE 1.2: Tree-level feynman diagrams of vector boson scattering. The amplitude of the diagrams without the Higgs boson, (a)-(c), are cancelled by the amplitude of the diagrams that involve the Higgs boson, (d), (e).

TABLE 1.2: Requirements on the objects used for the VBS measurements in the same-sign WW channel.

Requirement	ATLAS	CMS
Lepton (ℓ) requirements		
p_{T}^{ℓ} > [GeV]	27	25 (1st lepton) /20 (2nd lepton)
$\lvert\eta^{\ell}\rvert$ <	2.5	2.5
Jet (j) requirements		
p_{T}^{j} > [GeV]	65 (1st jet)/35 (2nd jet)	30
$\lvert\eta^{j}\rvert$ <	4.5	5.0
Requirements on the di-jet system (jj)		
m_{jj} > [GeV]	500	500
$\lvert\Delta\eta_{jj}\rvert$ >	2.0	2.5
Requirement on Missing E_{T} ($E_{\mathrm{T}}^{\mathrm{miss}}$, which is used to reconstruct neutrinos)		
$E_{\mathrm{T}}^{\mathrm{miss}}$ > [GeV]	30	40

1.2.2 Composite Higgs model

In a general model with a composite Higgs, we postulate the existence of a strongly interacting sector, apart from the elementary sector, which incorporates Higgs as a bound state and mixes with the elementary sector. In this model, the weak interaction is corrected by the usual Higgs mechanism without any fundamental scalars in the theory. A typical example of strongly coupled theories beyond the SM is 'Technicolor' [26]. In technicolor theories, inspired by chiral symmetry breaking in QCD, the electro-weak symmetry breaking is triggered by strong dynamics. However these theories are highly constrained by precision measurements of the electroweak interactions. Furthermore, incorporating flavour into technicolor theories is difficult [27].

One of the big differences of the composite Higgs models from technicolor is the existence of two separate phase transitions in the theory. We first postulate the existence of some "ultrafermions", at high energy regime, which interact via a strong "ultracolor" (UC) force, as well as the relatively weak SU(2) × U(1). Unlike in technicolor theories, the ultrafermions transform as a real representation of G_{UC}× SU(2) × U(1), where G_{UC} is the UC gauge group. Thus when UC force gets strong and causes an ultrafermion condensate (transition from a regime of free fermions to one of light bound states, including the Higgs doublet) to form, it does not have to *a priori* break SU(2) × U(1); this is followed by a second transition at the usual weak scale where the Higgs develops a Vacuum Expectation Value (VEV) and breaks SU(2) × U(l). An interesting and still economical variation is to have a Higgs arising as a composite pseudo-Goldstone boson (PGB) from the strongly interacting sector. In this case, the Higgs mass is "protected" by an approximate global symmetry and is only generated via quantum effects.

1.2.3 Benchmark model; Heavy Vector Triplets

Many BSM theories predict heavy resonances decaying into a pair of gauge bosons. It is *a priori* not possible to determine a "true" model to compare with data, nor practical to provide exclusion limits to any theories. Thus, we decided to adopt a model-independent analysis based on a simplified phenomenological Lagrangian; the Heavy Vector Triplet (HVT) model, which reproduces a large class of the BSM theories [28, 29].

The HVT model introduces a real vector triplet $V^a_\mu, a = 1, 2, 3$ $(W'^{\pm}, Z')$ in the adjoint representation of $SU(2)_L$ with vanishing hypercharge, as described in Equation 1.1, and the dynamics of the new vectors are described by the following Lagrangian in Equation 1.2:

$$V^{\pm}_\mu = \frac{V^1_\mu \mp iV^2_\mu}{\sqrt{2}}, V^0_\mu = V^3_\mu \quad , \tag{1.1}$$

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}D_{[\mu}V^a_{\nu]}D^{[\mu}V^{\nu]a} + \frac{m^2_V}{2}V^a_\mu V^{\mu a} \\ &+ ig_V c_H V^a_\mu H^\dagger \tau^a \overleftrightarrow{D}^\mu H + \frac{g^2}{g_V}c_F V^a_\mu J^{\mu a}_F \\ &+ \frac{g_V}{2}c_{VVV}\epsilon_{abc}V^a_\mu V^a_\mu V^b_\nu D^{[\mu}V^{\nu]} + g^2_V c_{VVHH}V^a_\mu V^{\mu a}H^\dagger H - \frac{g}{2}c_{VVW}\epsilon_{abc}W^{\mu\nu a}V^b_\mu V^c_\nu \quad .\end{aligned} \tag{1.2}$$

The first line of Equation 1.2 represent the V mass and kinetic energy, and the tri-linear and quadri-linear interactions with the SM gauge bosons from the covariant derivative:

$$D_{[\mu}V^a_{\nu]} = D_\mu V^a_\nu - D_\nu V^a_\mu, \quad D_\mu V^a_\nu = \partial_\mu V^a_\nu + g\epsilon^{abc}W^b_\mu V^c_\nu \quad , \tag{1.3}$$

where g denotes the $SU(2)_L$ gauge coupling. The second line in Equation 1.2 contains direct interactions of V with the Higgs current; g_V represents the coupling strength of the known W and Z bosons to the new vector bosons.:

$$iH^\dagger \tau^a \overleftrightarrow{D}^\mu H = iH^\dagger \tau^a D^\mu H - iD^\mu H^\dagger \tau^a H \quad , \tag{1.4}$$

and with the SM left-handed fermionic currents with constant C_F:

$$J^{\mu a}_F = \Sigma_f \bar{f}_L \gamma^\mu \tau^a f_L \quad , \tag{1.5}$$

where $\tau^a = \sigma^a/2$. The Higgs term and its coupling C_H leads to vertices involving the physical Higgs field and the three unphysical Goldstone bosons. In the high-energy region, the Goldstone bosons represent the longitudinal

polarisation of the W and Z bosons, thus C_H plays a significant role in both the production cross-section of the V resonance via Vector Boson Fusion (VBF) and its decay in the diboson mode.

Finally the third line of Equation 1.2 contains three new operators and free parameters (c_{VVV}, c_{VVHH} and c_{VVW}). However these terms in the third line do not directly contribute to V decays and single production processes that are only relevant for the analyses performed in LHC. They affect the processes through the mixing of V with W, however the mixing is typically small. Thus, as a first approximation, these terms can be ignored.

The $V^{\pm}$ or V^0 field is not a mass eigenstate and their mass arises from mixing with the SM gauge bosons and from the electro-weak symmetry breaking, therefore the mass term does not represent the actual mass of the resonances. We obtain a generalised relation between the masses of charged (M_+) and the neutral (M_0) resonance state [28]:

$$m_W^2 M_+^2 = \cos^2\theta_W m_Z^2 M_0^2 \quad . \tag{1.6}$$

In the high-mass limit ($M_{+,0} \gg m_{W,Z}$), considering the tree level $m_W^2/m_Z^2 \simeq \cos^2\theta_W$, Equation 1.6 implies the neutral and charged states will be degenerate in mass at a per-cent level.

As a consequence of the hierarchy assumption ($M_{+,0} \gg m_{W,Z}$), we derive a small mixing angle between the SM gauge bosons and the V states arising from the diagnolization of the matrices mixing the neutral and charged V fields with the $W^{\pm}, Z$ particles As all direct couplings of the V with W and Z bosons originate from mixing in the Lagrangian, a small mixing angle suppresses the decay width into vector bosons. With this effect, the decay width can be written as:

$$\Gamma_{V_0 \to WW} \simeq \Gamma_{V_\pm \to WZ} \simeq \frac{g_V^2 c_H^2 M_V}{192\pi} \quad . \tag{1.7}$$

As the V resonance decays into a pair of fermions or a pair of gauge bosons, the resonance can be produced in a collider through the quark-anti quark annihilation, gluon-gluon fusion, and the vector boson fusion processes [28]. In all these processes, the production cross-section can be written as [28]:

$$\sigma(pp \to V + X) = \Sigma_{i,j\epsilon p} \frac{\Gamma_{V\to ij}}{M_V} \frac{16\pi^2(2J+1)}{(2S_i+1)(2S_j+1)} \frac{C}{C_i C_j} \frac{dL_{ij}}{d\hat{s}}\bigg|_{\hat{s}=M_V^2} \quad , \tag{1.8}$$

where i, j denotes the particles interacting to produce the V resonance p=q, $\bar{q}$, W, Z, $\Gamma_{V\to ij}$, $S_{i,j}$ and $C_{i,j}$ the partial decay width into the particles, their spin and colour states. J and S account for the spin and colour stats of the final resonance and $\frac{dL_{ij}}{d\hat{s}}$ for the parton distribution function of the two initial state particles.

1.2.4 Extra Dimensions

One of the most fascinating scenarios in TeV-scale physics is that there are extra dimensions that allow us to solve the hierarchy problem by embedding the four-dimensional spacetime in a larger dimensional bulk with one single warped extra dimension. In this thesis, we consider a solution based on the Randall-Sundrum (RS1) framework [30]. A generic feature of this scenarios is the presence of spin-2 Kaluza-Klein (KK) excitations of the gravitons, whose masses and couplings to the SM particles are set in the TeV-scale. The KK graviton can appear as widely separated resonances, while we focus on the diboson resonances here.

In the original RS1 model, the SM particles are only allowed to be contained in the four-dimensional spacetime whereas the gravitational interaction can propagate through the bulk. However, this model introduces large contributions to flavour-changing neutral current processes and observables in electroweak precision measurements which are in contradiction with their current limits and measurements. The RS1 model is then extended to allow as well the SM gauge bosons and fermions to propagate in the additional dimension and is referred to the bulk RS1 model. In this case, graviton production and decay via light fermion channels are highly suppressed and the decay into photons are negligible.

The framework is based on a slice of Anti-de-Sitter$_5$ (AdS$_5$) [31] and the spacetime metric of the bulk RS1 model depends on the coordinate of the extra dimension. The hierarchy problem is solved by introducing an exponential warp factor:

$$ds^2 = e^{-2kr_c\phi}\eta_{\mu\nu}dx^\mu dx^\nu + r_c^2 d\phi^2 \quad , \tag{1.9}$$

where $\eta_{\mu\nu}$ is the Minkowski metric, x^μ is the four-dimensional coordinate, k is an energy scale, ϕ is the coordinate of the extra dimension ($0 \le \phi \le \pi$) and r_c is the radius of the curvature of the warped extra dimension. The Higgs

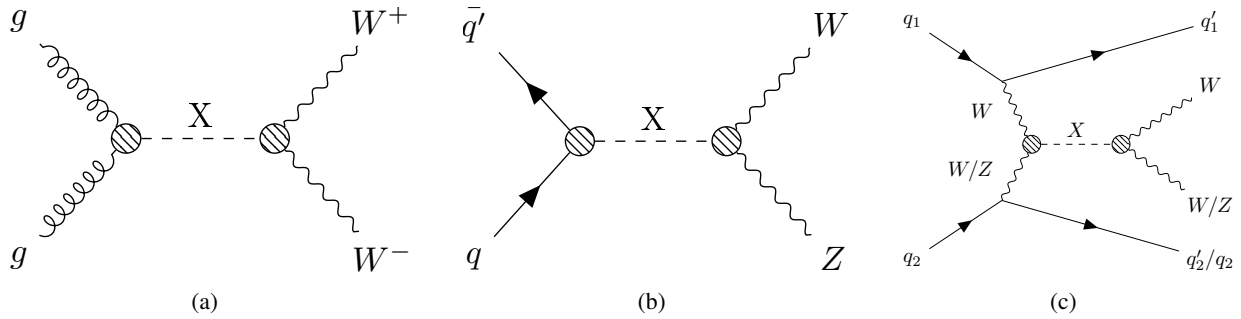


FIGURE 1.3: Production modes of the diboson resonances; gluon-gluon fusion (a), quark-antiquark annihilation (b) and vector boson fusion (c).

field is constrained to the "IR/TeV brane" with $\phi = \pi$ whereas the gravitational interaction is mostly localised at the "UV/Planckian brane" with $\phi = 0$. To generate this large hierarchy between the IR/TeV and UV/Planckian branes, no large radius of the warped extra dimension is needed ($kr_c \approx 11-12$) because of the exponential warp factor. The bench mark sample with the RS graviton decaying to WW is produced with $k/\bar{M}_{Pl} = 1.0$, where $\bar{M}_{Pl} = M_{Pl}/\sqrt{8\pi}$ is the effective four-dimensional Planck mass ($M_{Pl} = 2.4 \times 10^{18}$ GeV). In order to compare the results with the CMS one, we also perform an interpretation at $k/\bar{M}_{Pl} = 0.5$.[2] It is commonly believed that $k/\bar{M}_{Pl} \sim 1.0$ is outside of the range of validity of the model, however $k/\bar{M}_{Pl} \sim 1.0$ is still within the range of validity of the model [30]. Furthermore, based on the anti-de Sitter/conformal field theory (AdS/CFT) correspondence [33], the RS1 model is conjectured to be dual to the composite Higgs models.

Each SM particle propagating through the bulk results in KK expectations with masses at the TeV scale. For the massless graviton, the corresponding exited spin-2 KK gravitons are close to the TeV scale and thus their decay into fermions is suppressed and dominated by the decay into top-quark and Higgs boson pairs as well as WW and ZZ. The branching fractions of the WW and ZZ decay are approximately 18.7% to 16% and 9.5% to 8% respectively depending on the resonance mass and the width is about 6%.

1.3 Event topology of the signal processes

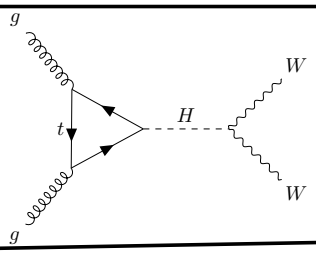
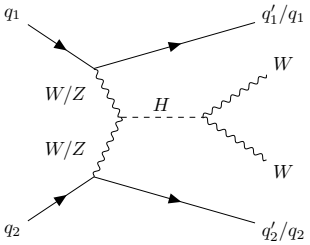
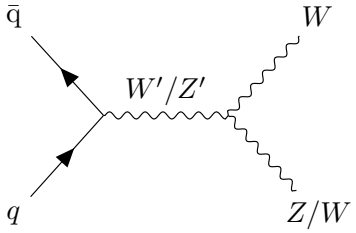
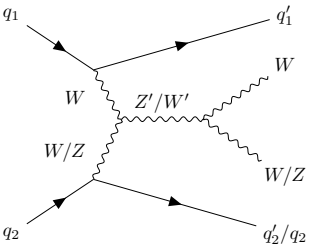
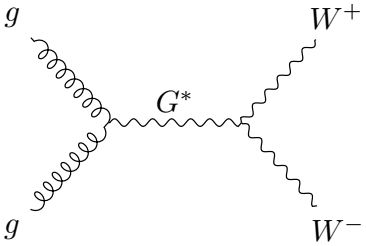
To maximise the sensitivity to the signal models, each signal region is separated into two orthogonal production-categories; gluon-gluon fusion or quark-antiquark annihilation (ggF/$q\bar{q}$) (Figure 1.3(a)/1.3(b)), and vector-boson fusion (VBF) (Figure 1.3(c)), depending on an assumed model.

The final state of the diboson resonances is characterised by the subsequent decays of the vector bosons. The searches in this thesis have been performed in the semileptonic final state, where one W boson decays leptonically and the other W/Z boson decays hadronically. For high-mass resonances, the opening angles between the quarks from the hadronically decaying V bosons are small because the V boson is highly-boosted; this effect is shown in Figure 1.4. The distance is defined as $R(q,\bar{q}) = \sqrt{\delta\eta^2 + \delta\phi^2} \sim (2m^V)/(p_{\mathrm{T}}^V)$. This case is referred to as the merged analysis where both quarks are included in a single jet, defined as a large-R jet, and this case is denoted by $\ell\nu J$, here J denotes a large-R jet. In contrast, separate identification of the two quarks from low-mass resonances is referred to as the resolved analysis and is denoted by $\ell\nu jj$, where j denotes a small-R jet.

The vector boson tagger described in Reference [35] is used to identify the hadronically decaying W/Z bosons in the merged analysis. In this tagger, large-R jets are identified as hadronically decaying W/Z bosons when they satisfies two requirements: (1) their mass is consistent with the W/Z boson mass and (2) the large-R jet has a sub-structure consistent with a two-prong jet. In the resolved analysis, the hadronically decaying V candidate is reconstructed by combining the two jets with the highest and second-highest p_{T} and requiring their invariant mass to be consistent with the W/Z boson mass. The object reconstruction schemes and the event selections are described minutely in Chapter 4 and 6 , respectively. Table 1.3 summarises the benchmark models used in this analysis.

[2] String theory favours $k/\bar{M}_{Pl} \sim 0.01$ [32]

TABLE 1.3: Summary of the benchmark models used in this analysis. The spin-0 interpretation is a neutral Higgs scalar boson decaying to two W boson. The heavy vector triplet model is used as a spin-1 interpretation. The spin-2 bulk Randall-Sundrum (RS) graviton is examined with values of $k/\bar{M}_{\mathrm{Pl}}$ = 1.0, 0.5

Model	Production	Diagram
Scalar Heavy Higgs Spin-0	ggF	
	VBF	
Heavy Vector Triplet (HVT) Spin-1	$q\bar{q}$	
Model-A: g_V=1 Model-B: g_V=3	VBF	
Bulk Randall – Sundrum (RS) Graviton Spin-2 $k/\bar{M}_{\mathrm{Pl}} = 1.0, 0.5$	ggF	

.

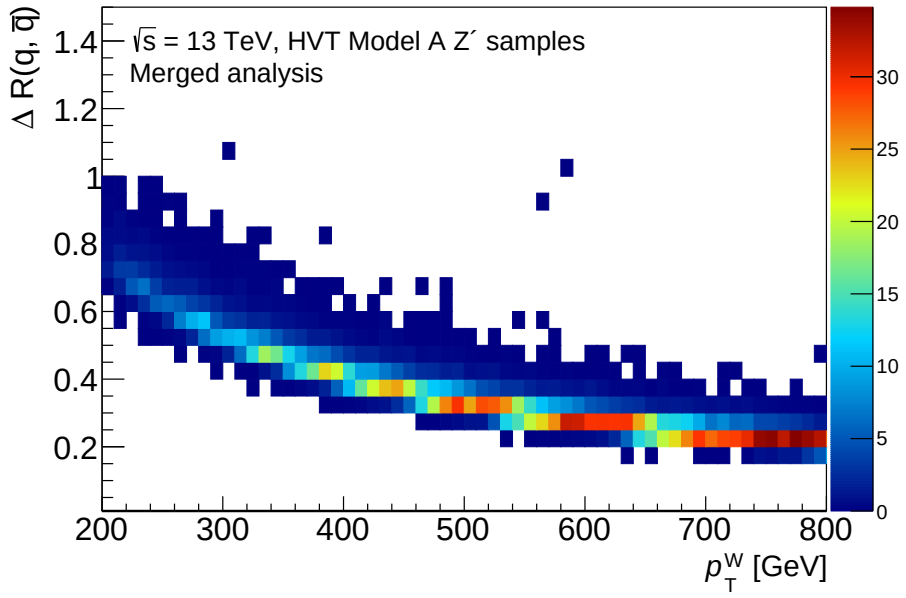


FIGURE 1.4: Angular distance between the quark and anti-quark, $\Delta R(q, \bar{q})$, from the hadronically decaying W boson produced in the HVT Z' samples as a function of the p_T of the W boson (p_T^W). We reconstruct quarks as jets using the anti-k_t algorithm [34] with a distance parameter $R = 0.4$. At the high p_T region ($p_T \gtrsim 400$ MeV), $\Delta R(q, \bar{q})$ is too small (i.e. $\Delta R(q, \bar{q}) < 0.4$) to reconstruct two quarks as two separate jets, thus we reconstruct hadronically decaying W/Z bosons as single large-R jets in the high p_T region.

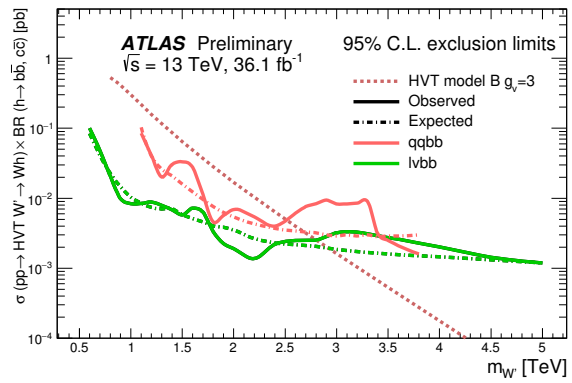
1.4 Current status of diboson resonance searches

The ATLAS and CMS collaborations have carried out diboson resonance searches in various decay modes. Table 1.4 summarises the results recently reported by the ATLAS and CMS collaboration using the 2015+2016 datasets, and Figure 1.5 shows the summary of the limits to the HVT W'/Z' cross-section.

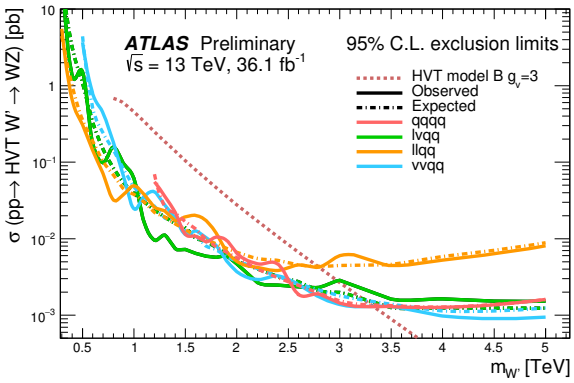
No significant excess in the events is yet observed with respect to the SM backgrounds. However, some small discrepancies are observed in several final states; e.g. deficit around 700 GeV in the $WZ \to \ell\ell qq$ final state observed both in the ATLAS and CMS results. To give a convincing explanation for them, larger integrated luminosity and smaller systematic uncertainties are essential.

TABLE 1.4: Current results of diboson resonance searches. The ATLAS and CMS collaborations have reported the results in various decay modes using 2015+2016 datasets.

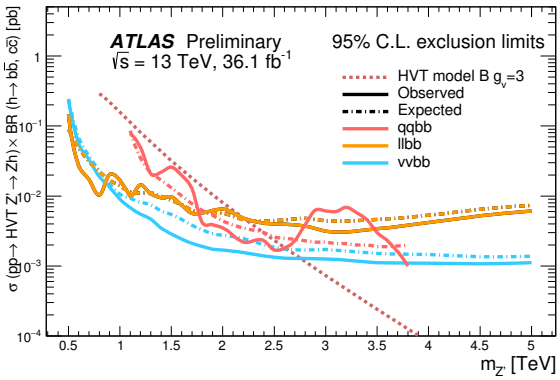
Collaboration	Category	Boson type	Final states
ATLAS	All hadronic	VV	$qqqq$ [36]
		VH	$qqbb$ [37]
	Semileptonic	WW/WZ	$\ell\nu qq$ [38]
		ZZ/WZ	$\ell\ell qq$ and $\nu\nu qq$ [39]
CMS	All hadronic	VV and qV	$qqqq$ and qqq [40]
		VH	$qqbb$ [41]
	Semileptonic	WW/WZ	$\ell\nu qq$ [42]
		ZZ/WZ	$\ell\ell qq$ [43]
		ZZ	$\nu\nu qq$ [44]
	Leptonic	ZZ	$\ell\ell\nu\nu$ [45]



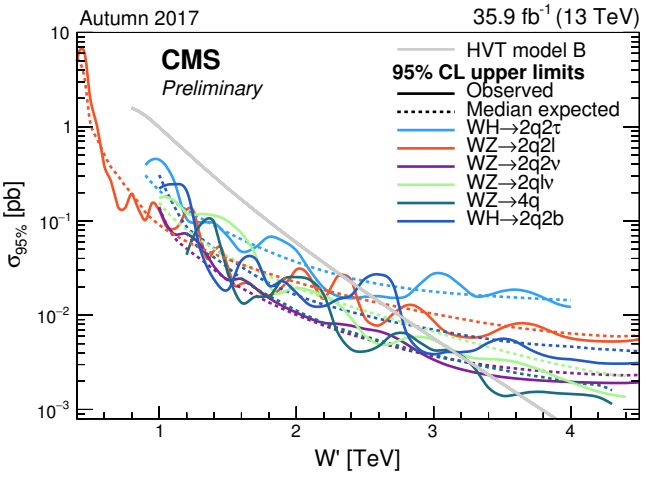
(a)



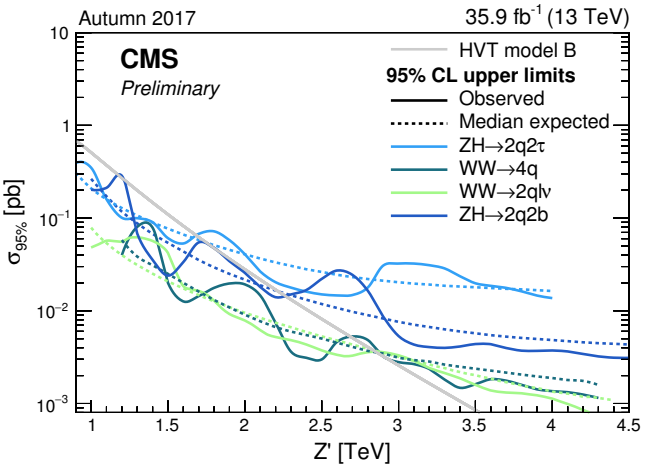
(b)



(c)



(d)



(e)

FIGURE 1.5: HVT W'/Z' limit summary, from the ATLAS collaboration (a)-(c) [46] and from the CMS collaboration (d), (e) [47]. The solid lines show the observed limits and the dashed lines show the expected limits. In the plots from the ATLAS (CMS) collaboration, brown dashed lines (gray solid lines) show the HVT model B cross-section which corresponds to a composite Higgs model. (a) shows W' limits for the final states of $W' \to Wh \to qqbb/\ell\nu bb$. (b) shows the W' limits for the final states of $W' \to WZ \to qqqq/\ell\nu qq/\ell\ell qq/\nu\nu qq$. (c) shows the Z' limits for the final states of $Z' \to ZH \to qqbb/\ell\ell bb/\nu\nu bb$. (d) shows W' limits for the final states of $W' \to WH \to 2q2\tau/2q2b$ and $W' \to WZ \to 2q2l/2q2\nu/2q\ell\nu/4q$. (e) shows Z' limits for the final states of $Z' \to ZH \to 2q2\tau/2q2b$ and $Z' \to WW \to 4q/2q\ell\nu$.

1.5 Structure of the thesis and Personal contribution

This thesis is structured as follows; for each chapter a brief introduction is given as well as the author contribution.

Chapter 1 gives an overview of the candidate BSM models which predict new particles decaying to a pair of vector boson.

Chapter 2 provides a general description of LHC and the ATLAS detectors.

Chapter 3 summarises the data and simulation samples used in the diboson resonance search.

Chapter 4 describes the algorithms and methods used in the analysis for reconstruction and identification of the particles produced in proton-proton collisions.

Chapter 5 reports the procedure and the results of the jet energy scale calibration for the large-*R* jets. I derived the jet energy calibration for the large-*R* jets, which is not yet used in the ATLAS public results but will be in the future results (Section 5.2). In chapter 10, we explain that the large-*R* jet related uncertainties is one of the largest uncertainty sources. In order to reduce the large-*R* jet uncertainties, data-driven "*in-situ*" measurements have been performed [48].

Chapter 6 provides a description of the event selection and its optimisation.

Chapter 7 reports a procedure for the background estimation.

Chapter 8 describes the sources of the systematic uncertainties.

Chapter 9 describes a general overview of statistical treatment in the diboson resonance search. I implemented the event selections described in chapter 6, and applied fitting methods on the events passed the selections. With applying the fit, various results have been obtained; post-fit distributions, impacts of the systematic uncertainties and expected/observed upper limits. Furthermore, the linear assumption to extrapolate the upper limits for the intermediate mass signals in the limit plots has been tested by a morphing technique considering a product of acceptance and efficiency of the neighbouring mass points. (chapter 10, B)

Chapter 10 presents the final results of the diboson resonance search.

Chapter 11 summarises the obtained results, then provides an outlook on possible measurements in the future.

Chapter 2

LHC and the ATLAS detector

The Large Hadron Collider (LHC) is a circular proton-proton collider currently running with a centre-of-mass energy of 13 TeV. The LHC beams are supplied from the injector chain (Linac, PS booster, PS, and SPS) and collide at four collision points, corresponding to the positions of four particle detectors; ATLAS, CMS, ALICE, and LHCb as shown in Figure 2.1. ATLAS and CMS are multi-purpose detectors to investigate new physics in the world highest energy scale, targetting measurements of a Higgs boson properties and searches for BSM. These two experiments are able to conduct basically same analyses, but have different detector designs. The two independently working collaborations enable cross-check of new results.

LHCb is a specialised detector to study CP violation processes through heavy flavour physics, and ALICE is an experiment optimised to study the heavy-ion collisions at LHC focused on physics of the quark-gluon plasma.

2.1 Large Hadron Collider

LHC was designed to collide protons (or heavy ions) with the centre-of-mass energy (E_{CM}) of 14 TeV (5.5 TeV/nucleon for heavy ion collision) [50]. LHC started the first running period (2010-2013, 'Run-1') with E_{CM} of 7-8 TeV [51]. To achieve higher energy (E_{CM} =13 TeV), LHC were shut down for the upgrades of the facilities since February 2013 (Long Shutdown 1, 'LS1'). LHC beam lines were improved during the LS1, and LHC has been successfully running at E_{CM} of 13 TeV in the second running period since 2015 (2015-2018, 'Run-2').

For a targetted physics that has a cross section of $\sigma_{\mathrm{physics}}$, a number of events generated in the collisions per unit time is given by:

$$N_{\mathrm{event}} = \mathcal{L}\sigma_{\mathrm{physics}} \quad , \tag{2.1}$$

where $\mathcal{L}$ is the instantaneous machine luminosity. When two Gaussian beams are colliding head-on, the luminosity $\mathcal{L}$ can be described as [50]:

$$\mathcal{L} = \frac{N_b^2 k_{\mathrm{B}} f_{\mathrm{rev}}}{4\pi\beta^* \epsilon_n} \times F \quad , \tag{2.2}$$

where N_b is the number of particles per bunch (protons for the LHC case), k_{B} is the number of bunches per beam, f_{rev} is the revolution frequency, β^* is the β-function at the collision point, ϵ_n is the normalised transverse beam emittance, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP), which must be smaller than 1 and is given by [50],

$$F = \left[1 + \left(\frac{\theta_{\mathrm{c}}\sigma_z}{2\sigma^*}\right)^2\right]^{-\frac{1}{2}} \quad , \tag{2.3}$$

where θ_{c} is the full crossing angle at the IP, σ_z the RMS bunch length and σ^* the transverse RMS beam size at the IP ($\sigma_z = \sqrt{\epsilon_n\beta^*}$). Table 2.1 summarises the LHC parameters in Run-1 and Run-2.

Equation 2.2 shows that there are several ways to increase the luminosity; decreasing β^* and ϵ_n, and increasing N_b and k_{B}. The luminosity in LHC reached $1.7 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ on 29th August, 2017, which is 1.7 times larger than the designed value. Figure 2.2(a) shows the delivered luminosity to ATLAS and Figure 2.2(b) shows the distribution of the mean number of interactions per bunch crossing. Figures 2.3 show delivered and recorded luminosity at ATLAS and the peak instantaneous luminosity in 2015, 2016 and 2017 runs.

In 2017, LHC adopted a new beam scheme; Batch Compression Merging and Splitting [55] (BCMS), which achieved a lower transverse beam emittance. While ramping-up the peak luminosity in the beginning of the 2017

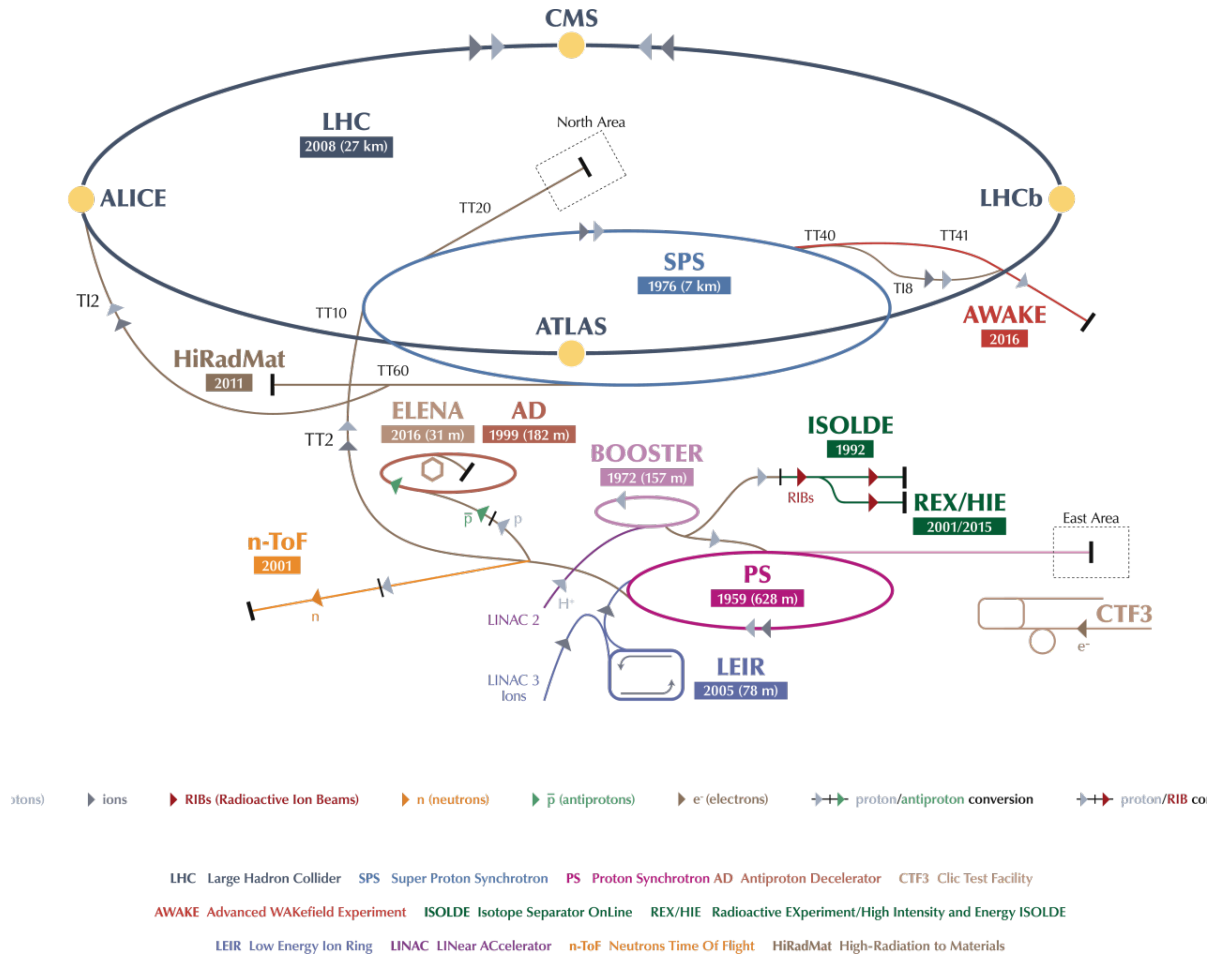
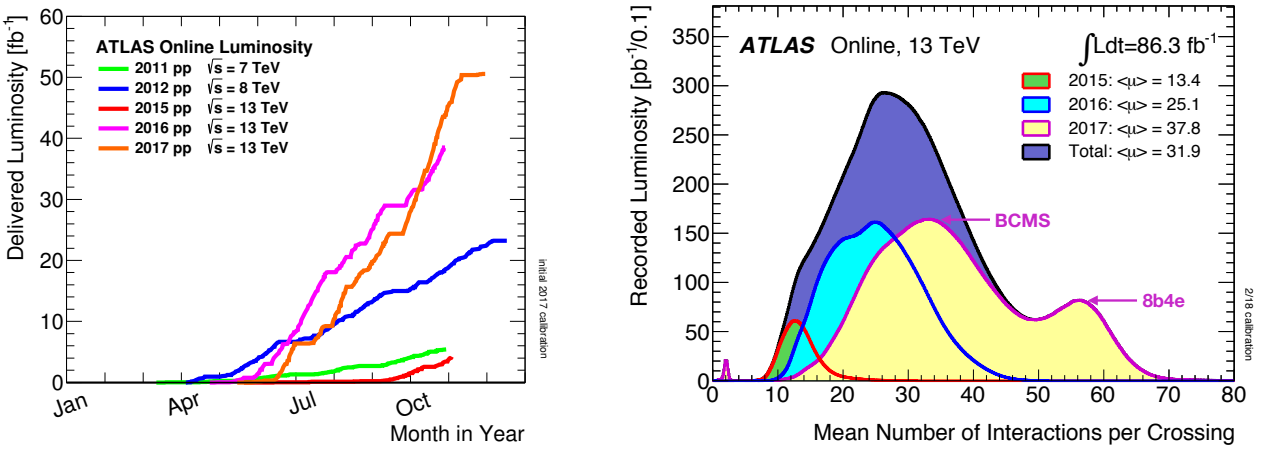


FIGURE 2.1: Schematic view of LHC and its injection chain at CERN. LHC is a circular proton-proton collider with a circumference of 27 km. [49]

TABLE 2.1: Important parameters of LHC [52], [53] and [54].

Parameter	Design	Run-1	Run-2			
		2010-2012	2015	2016	2017 BCMS	2017 8b4e
Centre-of-mass energy (E_{CM}) [TeV]	14	7-8	13	13	13	13
Bunch spacing ($1/f_{rev}$) [ns]	25	50	50	25	25	(25)
Normalised transverse beam emittance (ϵ_n) [mm × mrad]	3.75	2.5	3.5 (2.5)	2.3	2.2-2.3	(1.5-2.0)
Bunch population (N_b) [$10^{11}p$]	1.15	1.0-1.5	1.1	1.1	1.3	1.25
Number of bunches per beam (k_B)	2808	1380	2244	2200	2556	1868
Peak luminosity [$10^{34}cm^{-2}s^{-1}$]	1.0	0.77	0.51	1.38	1.7	(1.5)
Optimal β-function (β^*) [cm]	55	60	80	40	40	30
Averaged number of interactions per bunch crossing	~23	~20	13.4	25.1	~ 40	~ 60



(a) (b)

FIGURE 2.2: (a) Integrated luminosity delivered to ATLAS during the high energy pp collisions. Both Run-1 and Run-2 data are shown as a function of running time. (b) Mean number of interactions per bunch crossing for the 2015, 2016 and 2017 pp collision data at $\sqrt{s} = 13$ TeV [59].

run, a problematic cell "16L2" was found; the 16L2 cell caused beam losses. To avoid the 16L2 issue, the standard LHC beam scheme with a bunch spacing of 25 nanoseconds was replaced by a new beam scheme; "8 bunches and 4 empty buckets" (8b4e) [56]. The 8b4e beam consists of mini-trains of eight consecutive bunches spaced by 25 nanoseconds and four empty bunch slots. LHC had also started to exploit a new optics, namely the "achromatic telescopic squeezing" (ATS) scheme [57]. which allowed us to reduce the β^* from 40 cm to 30 cm, increasing the luminosity[1]. However, higher instantaneous luminosity ($\sim 2.0 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$) is achievable with a smaller number of bunches per beam (1868) by introducing the 8b4e scheme, hence the number of interactions per crossing increased. Thus, we introduced a luminosity levelling, namely separation levelling [58], to keep the instantaneous luminosity at $1.55 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ keeping the number of interactions per crossing at ~ 60.

Thanks to the improved LHC performance, scientists can exploit the higher statistics, however, as shown in Figure 2.2(b), more than 60 interactions per crossing are observed in the latter part of the 2017 run. It is challenging to operate the detectors and the trigger system in this higher pile-up condition keeping the reconstruction efficiency for target physics objects.

2.2 ATLAS detector

In this section, we overview the current performance of the ATLAS detector. The ATLAS detector [60] at LHC is a general purpose detector designed for the measurements of the SM parameters (especially of the Higgs boson and the top quark) and searches for BSM. It consists of three major types of detector systems; the inner detector system, the calorimeter system, and the muon spectrometer. All the information from each component is used for the object reconstruction and/or the triggering system. Figure 2.4 shows a schematic view of the ATLAS detector.

We employ the cylindrical coordinate system (r, z, ϕ), where the z-axis is defined along the beam pipe direction. The rapidity y is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \quad , \tag{2.4}$$

where E and p_z are the energy and z-component of the momentum of a given particle. The pseudo rapidity η, which is the massless limit of the rapidity, is defined as follows:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \quad . \tag{2.5}$$

[1] Peak luminosity $2.0 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ was achieved in a test run

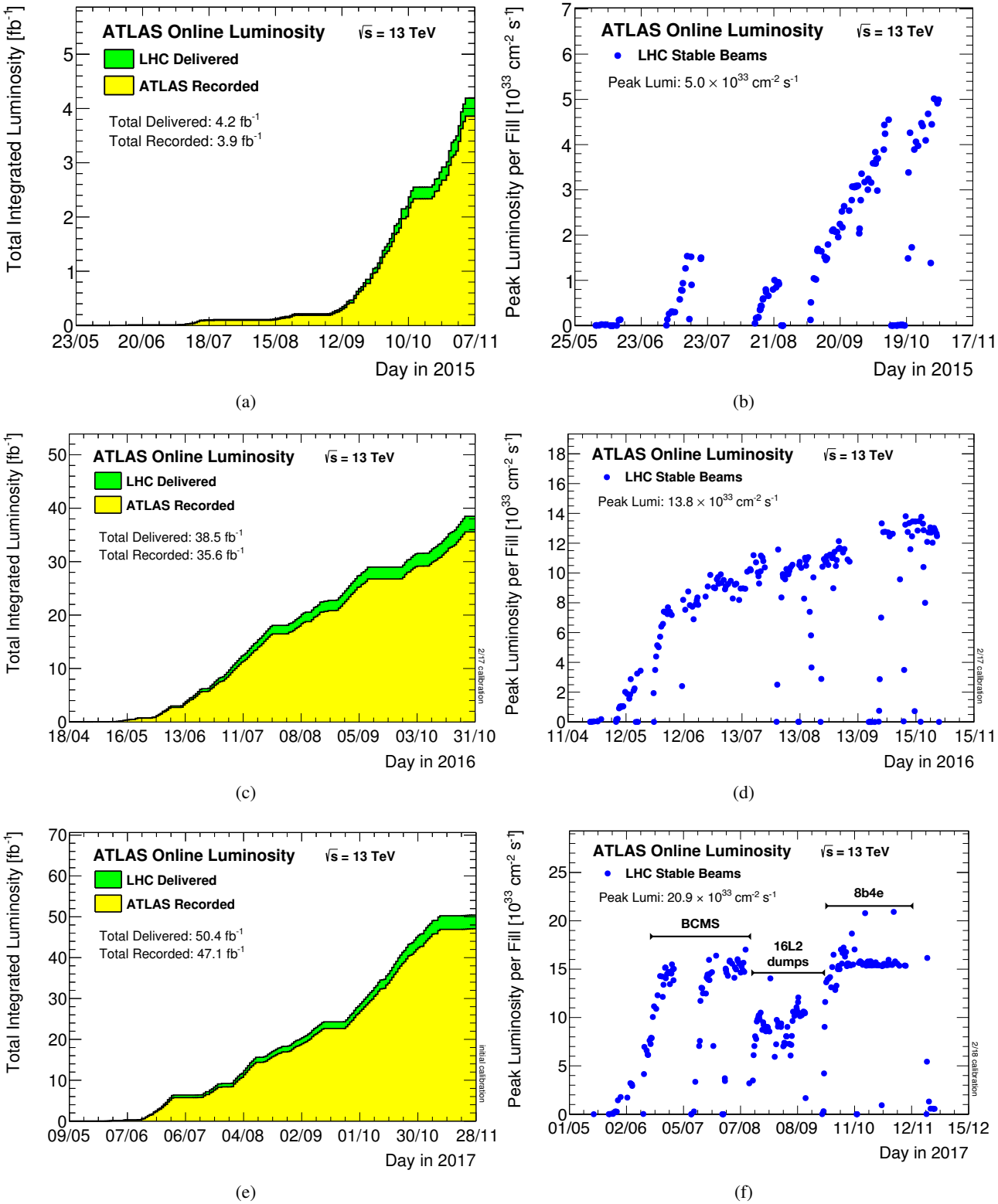


FIGURE 2.3: (a), (c) and (e) the cumulative luminosity versus time delivered to (green) and recorded by ATLAS (yellow) during pp collisions at 13 TeV centre-of-mass energy in 2015, 2016 and 2017, respectively. (b), (d) and (f) the peak instantaneous luminosity delivered to ATLAS during pp collisions at $\sqrt{s} = 13$ TeV for each LHC fills as a function of time in 2015, 2016 and 2017, respectively. [59]

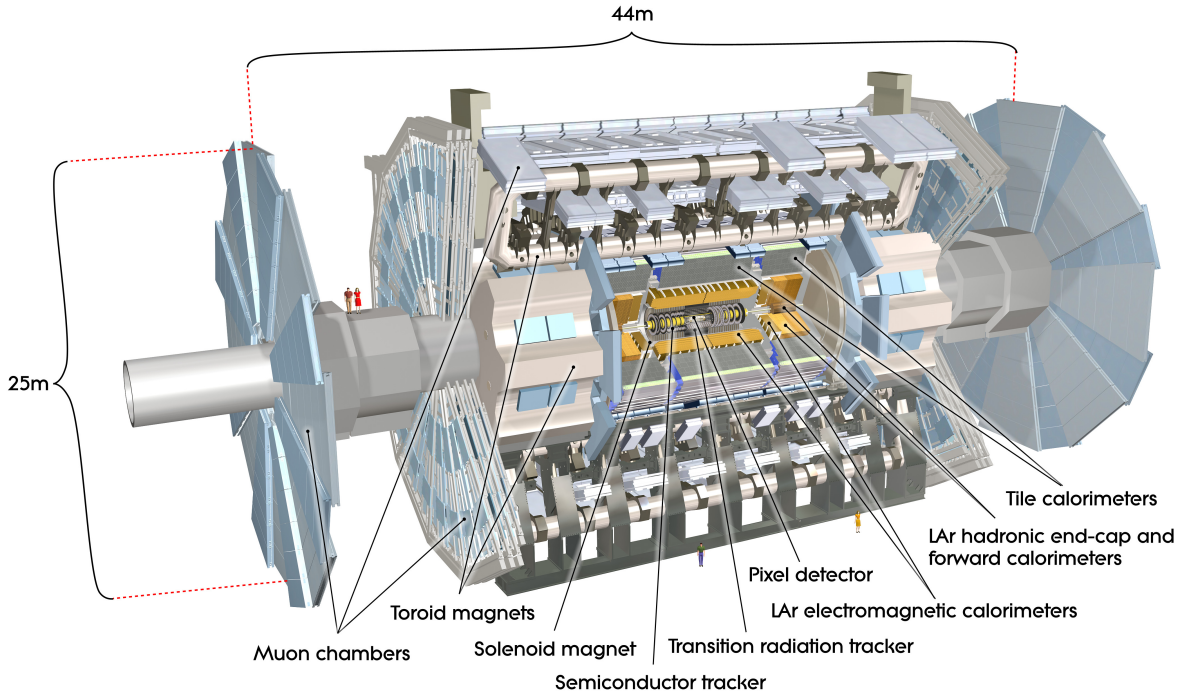


FIGURE 2.4: Schematic view of the ATLAS detector consisting of three parts of detectors; the inner detector, the calorimeter and the muon spectrometer [60].

A schematic view of the ATLAS coordinate system is shown in Figure 2.5. The x-y plane is perpendicular to the beam line and is referred to as the transverse plane. η is preferred rather than the polar angle because particle production is approximately constant as a function of η in the hadron collider physics.

There are three main components of the inner detector system (ID); a high-granularity silicon pixel detector including a newly installed innermost layer called the insertable B-layer (IBL) [61], a silicon microstrip tracker (Semiconductor tracker, SCT) and a straw-tube tracker (Transition radiation tracker, TRT). ID is surrounded by a superconducting solenoid magnet, which provides a 2 T magnetic field which enables ID to provide charged particle tracking and momentum measurement in the range of $|\eta| < 2.5$. The calorimeter system consists of electromagnetic and hadronic calorimeters covering the range of $|\eta| < 4.9$. The precision electromagnetic calorimeters are composed of liquid argon (LAr) with accordion-shaped absorbers and electrodes. The hadronic calorimeters consist of iron plates and plastic scintillator tiles, and LAr as well. The muon spectrometer system is composed of two types of detectors; trigger and high-precision measurement chambers, measuring deflection of muons in a superconducting toroidal magnetic field to calculate momentum of muons. The muon trigger system is composed of resistive plate chamber (RPC) for the barrel region ($|\eta| < 1.05$) and thin gap chamber (TGC) for the endcap region ($1.05 < |\eta| < 2.4$). The precision chamber system covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes (MDT), complemented by cathode strip chambers (CSC) in the forward region, where the background is highest.

2.2.1 Inner detector

A schematic view of the inner detector system is shown in Figures 2.6.

Pixel detector

The pixel detector consists of four barrel layers centred around the beam axis, and three disks on each side. The main purpose of the pixel detector is to measure the charged particle trajectories to reconstruct primary and secondary vertices precisely, with a coverage of $|\eta| < 2.5$. The standard pixel size is 50×400 μm^2. The position resolution is 10 μm in the R-ϕ plane and 115 μm in z (R) for the central (endcap) region. We achieve this high-resolution measurement using pulse-height information [63].

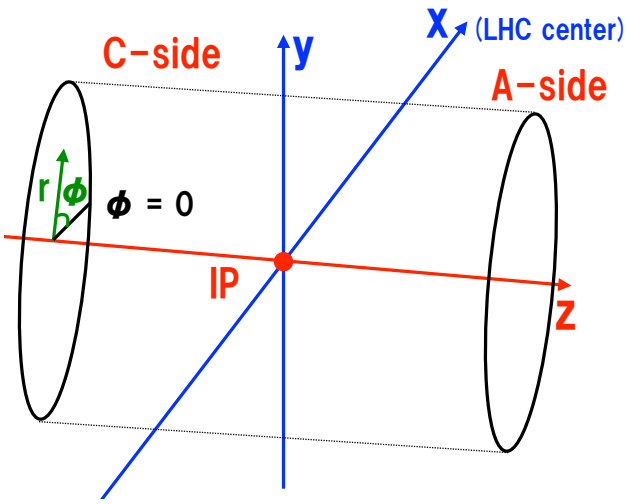
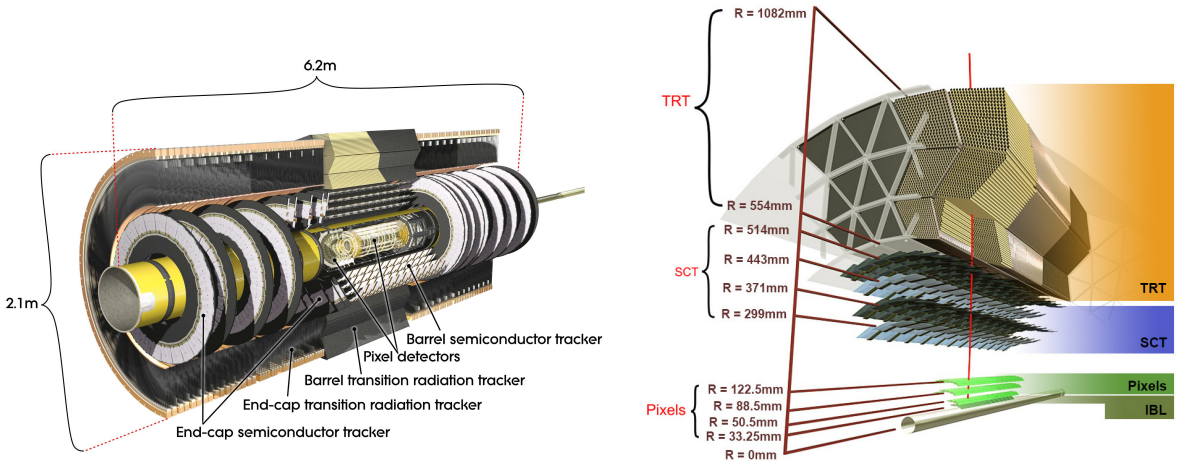


FIGURE 2.5: Schematic view of the ATLAS coordinate system. We use the cylindrical coordinate taking the z-axis along the beam pipe.The positive z region is referred to as the A-side, and the other side is referred to as the C-side.



(a) (b)

FIGURE 2.6: Schematic view of (a) the inner detector system [60], with (b) a detailed layout of the barrel region [62]. The inner detector system is composed of three detectors: high-granularity silicon pixel detectors (Pixel) closest to the interaction point, the silicon strip semiconductor Tracker (SCT), and the outermost detector, transition radiation tracker (TRT).

The innermost layer, the insertable B-Layer (IBL), is located between the new thinner beam-pipe and the second innermost Pixel layer, the B-Layer, which is the innermost layer during the LHC Run-1. As described in Section 2.1, the peak luminosity has reached 1.7×10^{34} cm^{-2}s^{-1} in the LHC Run-2, however the B-Layer was designed to deal with 1×10^{34} cm^{-2}s^{-1}. To compensate the B-Layer degradation due to irreversible radiation damage, IBL was installed in the LHC Run-2.

Silicon strip semiconductor tracker

The semi conductor tracker (SCT) is a silicon microstrip detector, and is the second part of the inner detector covering the region of $|\eta| < 2.5$. The barrel region of SCT consists of eight strip layers, each of which is composed of two 6.4 cm daisy-chained sensors with a strip pitch of 80 μm. In the endcap region, nine discs are installed on each side, using two radial layers of strips each. The SCT modules are constructed from two pair of micro-strip sensors (doublet) with a tilt angle of 40 mrad. Each of the doublet layers achieves a two-dimensional measurement. SCT is designed such that each particle within its coverage traverses four doublet layers. The spatial resolution of SCT is 17 μm in the $R - \phi$ plane and 580 μm in the Z (R) for the barrel (endcap) region.

Transition radiation tracker

The transition radiation tracker (TRT) is the third and outermost part of the inner detector system. The TRT hits are used for tracking as well as for electron identification. TRT is composed of straw tubes which are 4 mm in diameter filled with a mixture of gases 70% Xe, 26% CO_2 and 3% O_2. In the barrel region, the tubes are 144 cm long and aligned parallel to the beam-pipe, thus TRT provides no information on the z-position whereas the resolution on the (R, ϕ) is about 130 μm. In the endcap region, the 37 cm long straws are arranged radially in wheels.

The electron identification is performed using the transition radiation X-ray photons produced by relativistic electrons at the boundary between the straws and the gas. The intensity of the transition radiation is proportional to the Lorentz factor $\gamma = E/m$. Using the Lorentz factor, electrons can be identified from hadrons whose momenta are same with the electrons; the mass of electrons can be ignored at the high energy region and thus its transition radiation is above a characteristic threshold, on the other hand, the intensity of radiation for heavy objects like hadrons is much lower.

2.2.2 Calorimeter

The calorimeters locate outside of the solenoid magnet. The calorimeter system, which covers the range of $|\eta| < 4.9$, is composed of the Liquid Argon (LAr) electromagnetic calorimeter, the Tile calorimeter, the LAr hadronic end-cap calorimeter, and the LAr forward calorimeter as shown in Figure 2.7. The calorimeter system plays a crucial role to reconstruct the electron/photon and hadronic jets in the diboson resonance searches as well. Furthermore, we exploit the calorimeter information to reconstruct neutrinos or BSM weakly interacting neutral particles with the missing transverse energy ($\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\Sigma_{\text{physics objects}} \vec{p}_{\mathrm{T}}$) in the event.

Liquid argon electromagnetic calorimeter

The electromagnetic calorimeter is composed of a barrel region ($|\eta| < 1.475$) and two end-cap parts ($1.375 < |\eta| < 3.2$), each of which is a lead-LAr sampling calorimeter with accordion-shaped absorbers and kapton electrodes. The accordion structure covers possible azimuthal cracks and provides complete ϕ symmetry. The innermost layer has the finest granularity along η so that, for example, we can distinguish individual photons from those arising from π^0 decays, examining the shower shape at its earliest stage. The total thickness is ~22 X_0, where X_0 is the radiation length; the middle layer has a thickness of 16 X_0 and thus collects the bulk part of the energy. The outer layer has a much coarser granularity. The main purpose of the outer layer is to correct the overlap of the energy deposition in the following hadronic calorimeter.

The absorbers are made of lead plates which have a thickness of 1.53 mm for $|\eta| < 0.8$ and of 1.13 mm for $|\eta| < 0.8$. A fine granularity pre-sampler detector, which covers the range of $|\eta| < 1.8$, corrects energy loss upstream of the calorimeter. The structure described above is shown in Figures 2.8 and 2.9. Figure 2.8 shows the radiation length before and inside the electromagnetic calorimeter as a function of η. Figure 2.9 shows the interaction length in the electromagnetic and hadronic calorimeters as a function of η.

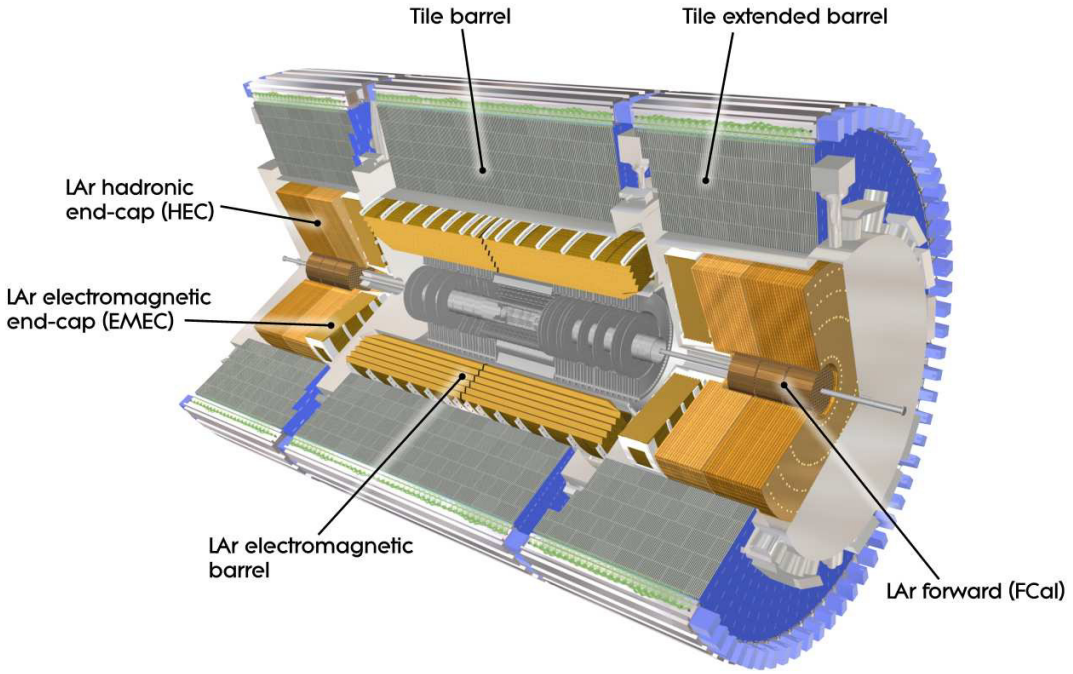


FIGURE 2.7: Schematic view of the calorimeter system composed of four types of detectors: the LAr electromagnetic calorimeter, the Tile calorimeter, the LAr hadronic end-cap calorimeter and the forward calorimeter. [60]

The resolution of the calorimeter can be described with a following expression [64]:

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

where a, b, and c are η-dependent parameters; a is the sampling term, b is the noise term, and c is the constant term. The sampling term is from fluctuation related to the stochastic development of showers, and thus contributes mostly at low energy compared to the constant term; its design value is (9-10)%/ $\sqrt{E[\mathrm{GeV}]}$ at low $|\eta|$. At larger η, this term is expected to be worse as the amount of material in front of the calorimeter is larger. The noise term comes from the electronic noise of the readout chain and depends on the detector technique and on the readout system. It is about $350 \times \cosh \eta$ MeV for a typical cluster in the barrel region and for a mean number of interactions per bunch crossing $\langle \mu \rangle = 20$. At larger η, this term is dominated by the pile-up noise. At higher energy, the sampling term and the noise term become less important and the relative energy resolution tends asymptotically to the constant term, c, which has a design value of 0.7 %. The constant term does not depend on the energy of the particle and is originating from the dead material in the calorimeter and the non-uniformity of the calibration of the calorimeter.

Tile calorimeter

The Tile calorimeter is a sampling calorimeter with steel and plastic scintillators (tile). It consists of a barrel region $|\eta| < 1.0$ and two extended barrel region $0.8 < |\eta| < 1.7$. Each of barrel and extended barrels is divided azimuthally into 64 modules, and is segmented in radial depth into three layers as shown in Figure 2.10. Each cell has the dimensions of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the first and the second layers and 0.1×0.2 for the third layer. Wavelength shifting fibres and photomultiplier tubes are connected to both sides of the scintillator.

There are some special cells at the gap region between the barrel and the extended barrels, to partially recover the missing energy in the gap region. The tile calorimeter has three calibration systems: charge injection, laser, and ^{137}Cs radioactive source. The photomultiplier gains and the front-end electronics are calibrated combining the three calibration systems. The readout signals are transformed to a unipolar pulse shape and amplified. The fractional energy resolution, which is parametrized by Equation 2.6, is measured by the pion test-beam, and the parameters a, b are measured to be 56 and 5.5 %, respectively [60].

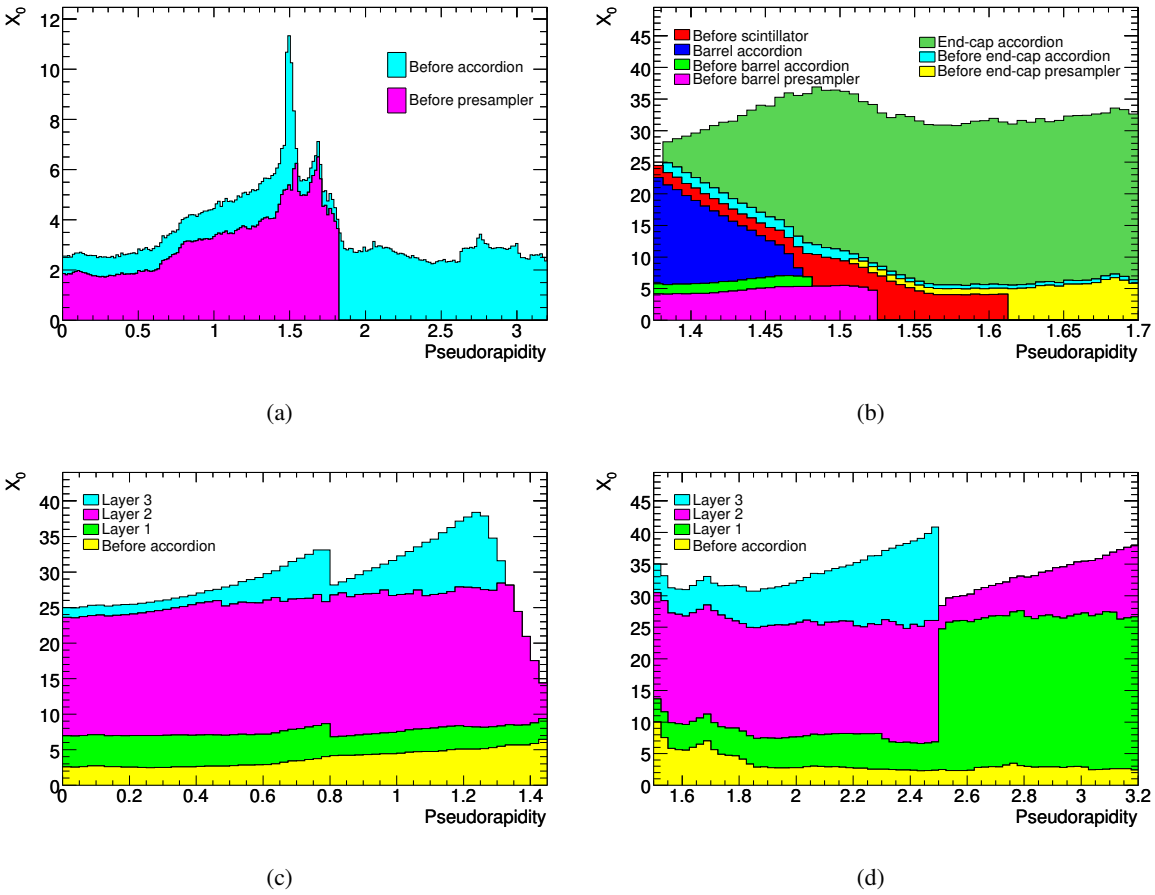


FIGURE 2.8: Cumulative amount of materials in front of and in the electromagnetic calorimeter in units of radiation length X_0 as a function of $|\eta|$ [60]. (a) the total amount of material upstream of the calorimeter. (b)-(d) show the total amount of material in the calorimeter. (b) the detailed view of the region between the barrel and end-cap regions. (c) and (d) show the barrel and end-cap regions, with three layers of the electromagnetic calorimeters in the range of $0 < |\eta| < 2.5$ or two layers in the higher-η region $2.5 < |\eta| < 3.2$.

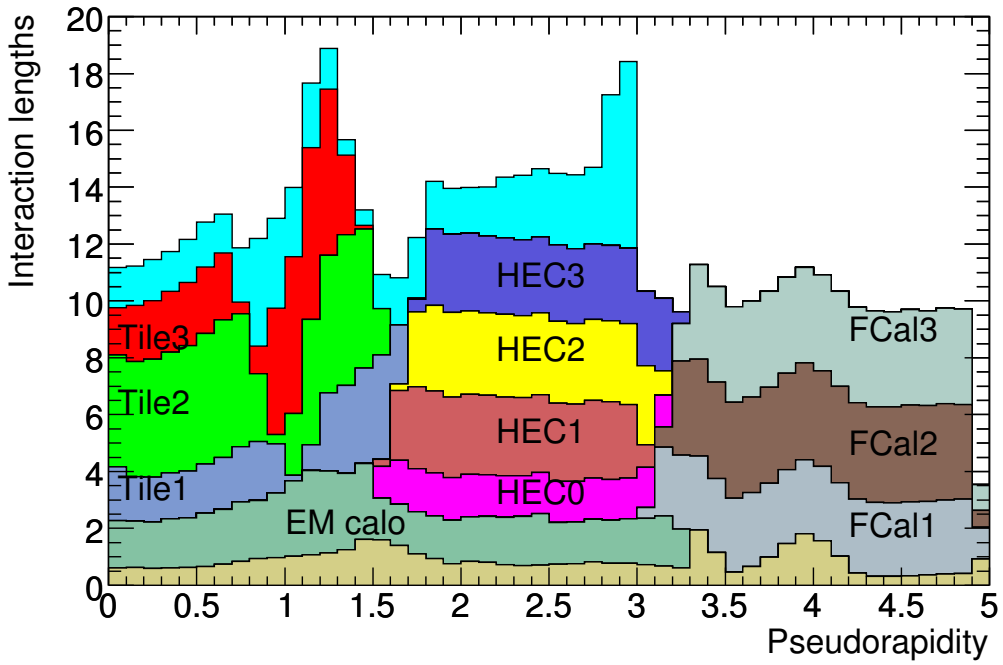


FIGURE 2.9: Cumulative amount of material, in units of interaction length, as a function of η at the calorimeter, in the electromagnetic calorimeter and in the hadronic calorimeters [60].

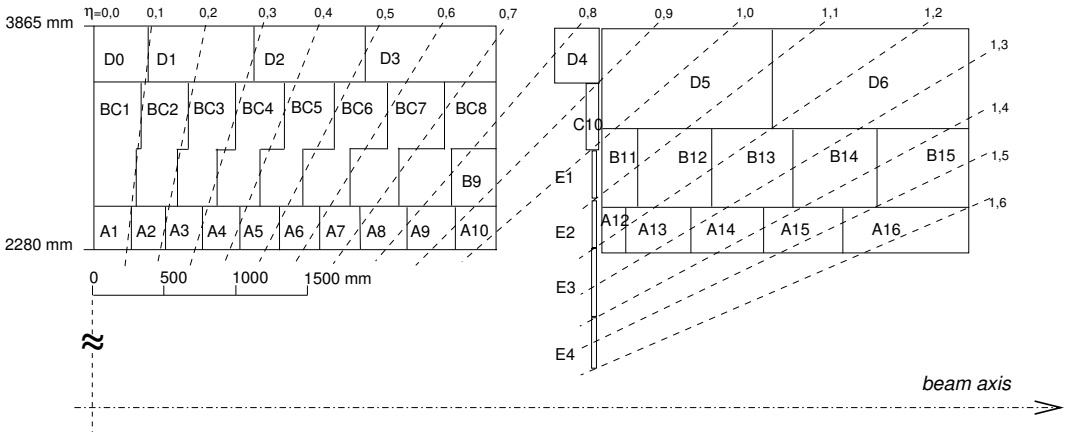


FIGURE 2.10: Segmentation in depth and η of the tile calorimeter in the central (left) and extended (right) barrels. The bottom of the picture (r = 2280 mm) corresponds to the inner radius of the tile calorimeter. [60].

TABLE 2.2: Resolution parameters of the four muon sub-systems. The spacial resolution does not include the alignment uncertainties and the time resolution does not account for the signal-propagation and electrons contributions [60].

		Chamber resolution (RMS) in		
Type	Function	z/R	ϕ	time
MDT	precision-tracking	35 μm (z)	—-	—-
CSC	precision-tracking	40 μm (R)	5 mm	7 ns
RPC	trigger	10 mm (z)	10 mm	1.5 ns
TGC	trigger	2-6 mm (R)	3-7 mm	4 ns

Liquid argon hadronic end-cap calorimeter (HEC)

The hadronic end-cap calorimeter (HEC) is a copper-LAr sampling calorimeter with a flat-plate design. It covers the end-cap region of $1.5 < |\eta| < 3.2$. HEC consists of two wheels, each of which consists of two longitudinal sections, thus it has four layers. Each of the HEC wheels is azimuthally segmented into 32 identical wedge-shaped modules.

The read-out granularity of HEC is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the range of $1.5 < |\eta| < 2,5$ and 0.2×0.2 for the range of $2.5 < |\eta| < 3.2$. The energy resolution of HEC is measured in the pion test-beam as well to be $a, b = 70.6, 5.8$ %, respectively. It is worse than the LAr calorimeter or TileCal because the energy deposited has greater fluctuations due both to the effect of the electromagnetic component of the jet and to the presence of energy that cannot be detected as excitation, fission energy or as energy lost by neutrino emission.

Forward calorimeter (FCal)

The Forward calorimeter (FCal) consists of three modules: the absorber in the first module is made of copper and is focused on electromagnetic measurements. In the other two modules, the absorber is made of tungsten. The gap between absorbers as an active medium of LAr is thin (0.25 mm) to minimise the effect of the radiation damage and to achieve fast readout. FCal has clear benefits in terms of resolution on the missing transverse energy and efficient tagging of forward jets, which are keys for the vector-boson fusion and vector-boson scattering processes. It needs to sustain a higher dose of radiation and to be operated in a higher rate environment in the forward region ($3.1 < |\eta| < 4.9$).

2.2.3 Muon spectrometer

The muon spectrometer forms the outermost part of the ATLAS detector in the region of $|\eta| < 2.7$. It is composed of two types of detectors: the trigger chambers and the precision-tracking chambers. A cross-section of the muon spectrometer is shown in Figure 2.11 and the spacial and timing resolutions of the four muon sub-detectors are summarised in Table 2.2.

The measurement of muon momentum is based on the curvature of the trajectories by the superconducting toroid magnets. Figure 2.12 shows the schematic view of the ATLAS magnet system. The toroidal magnetic field is generated by the barrel and end-cap toroid magnet systems. The cylindrical volume surrounding the calorimeters and the end-cap toroid magnets is immersed in the magnetic field of the barrel toroid magnet, which consists of eight superconducting coils. The coils are assembled in 8-fold symmetry in ϕ, and the overall size of the barrel toroid magnet system is 25 m in length, with the inner and outer diameters of 9.4 and 20 m, respectively. Each of end-cap toroid magnets is composed of a cold mass, which are mounted in a cryostat. The performance in terms of bending power is characterised by the field integral $\int \mathrm{B}dl$, where B is the field strength perpendicular to the muon direction. The integral is computed along an infinite-momentum muon trajectory between the innermost and the outermost muon-chamber planes.

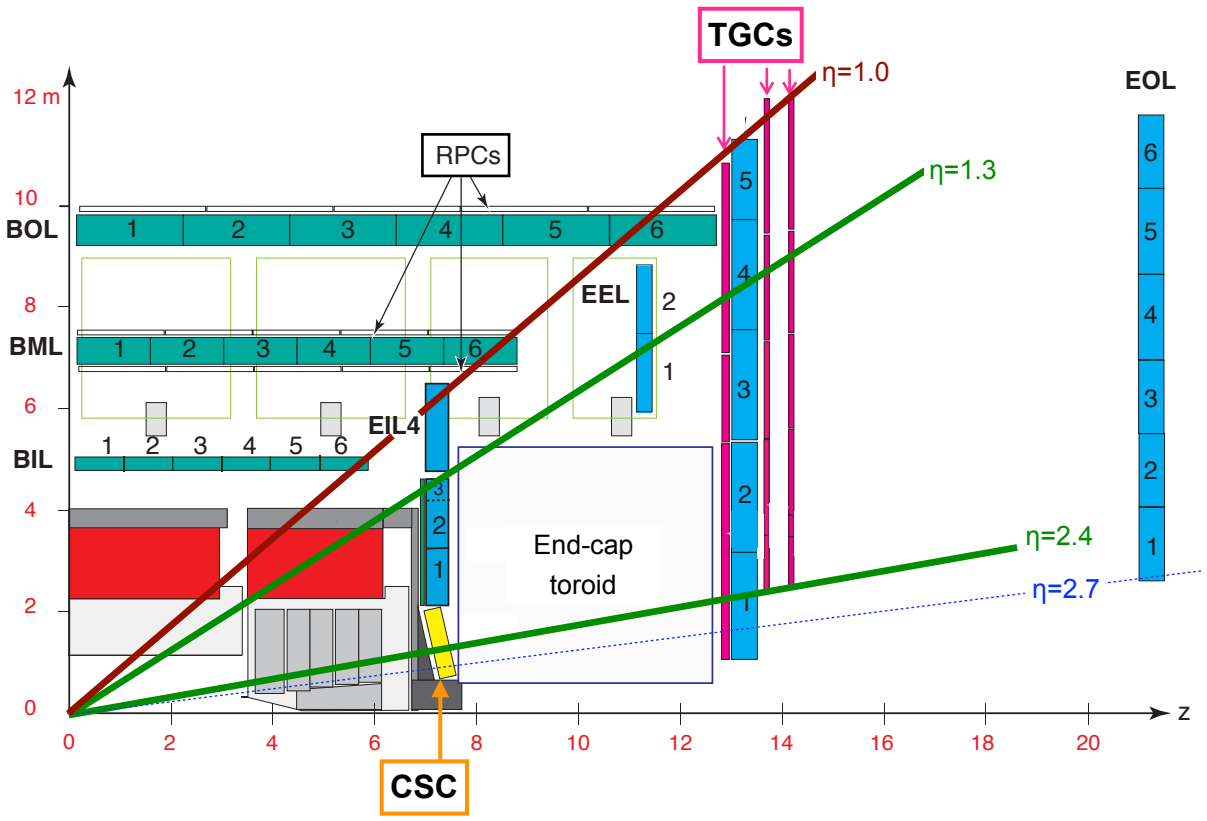
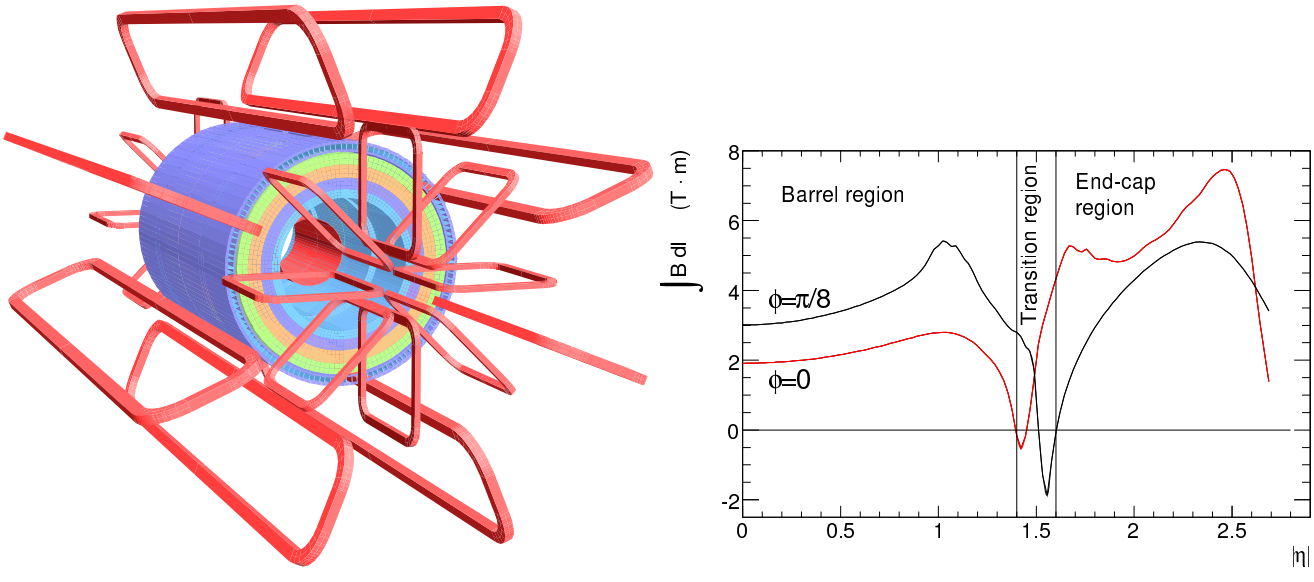


FIGURE 2.11: Cross-section of the Muon spectrometer. The muon spectrometer is composed of the two types of detectors: the muon trigger chambers (RPC for the Barrel region and TGC for the Endcap region) and the muon precision-tracking chambers (MDT and CSC) [60]. The blue (green) boxes indicate the Endcap (Barrel) MDTs. CSC is very close to the beam line, thus the rate of particles going through the chambers is high.



(a) (b)

FIGURE 2.12: (a) Schematic view of the magnet system, which consists of a solenoid magnet and toroidal magnet systems. The toroid magnet has 8-fold symmetry in ϕ. (b) The predicted filed integral in one toroid octant as a function of $|\eta|$. The toroidal magentic field is divided into two regions: the barrel and the end-cap regions. The field strength is not enough in the intermediate region of the barrel and endcap ($\eta \sim 1.4$ for $\phi = 0$, $\eta \sim 1.6$ for $\phi = \pi/m$) to measure p_{T} of the muons precisely [60].

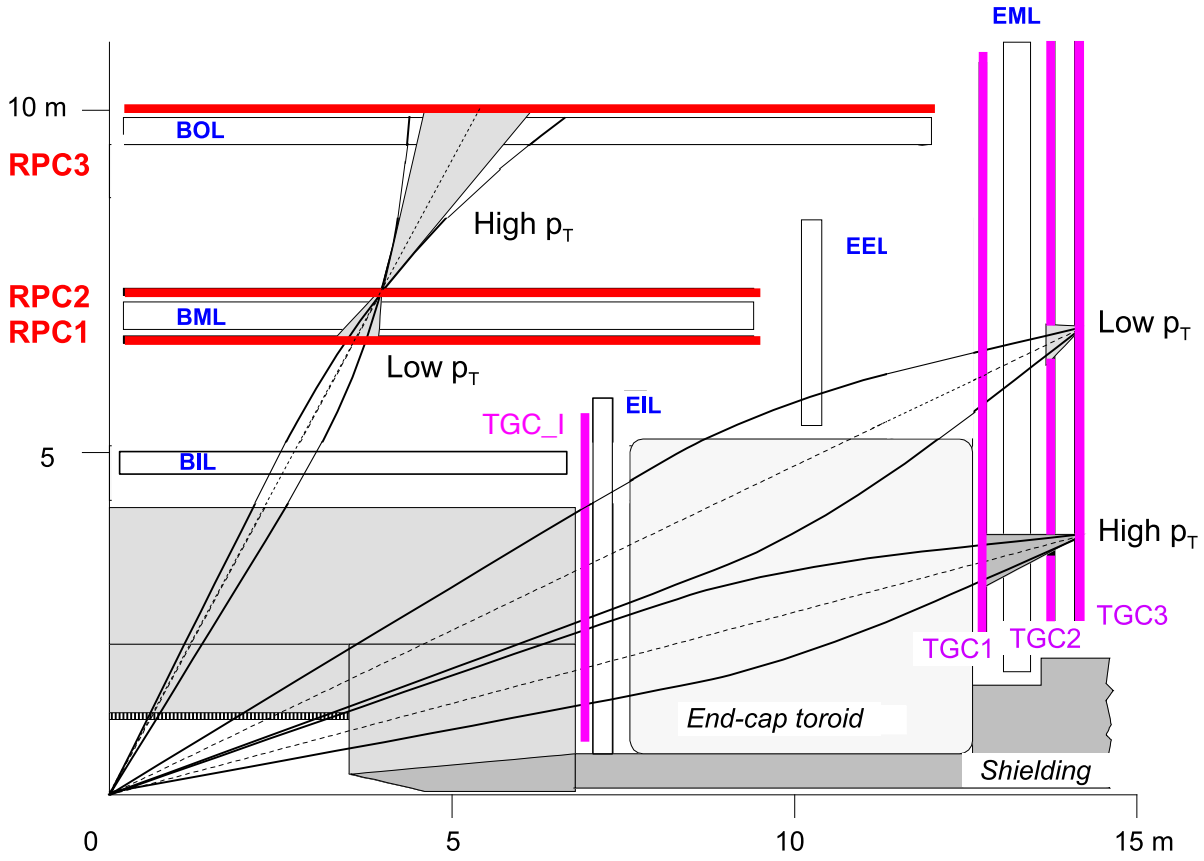


FIGURE 2.13: Schematic view of the muon trigger system. RPC2 and TGC3 are the reference (pivot) planes (i.e. define trigger positions) for barrel and end-cap regions, respectively [60]. TGC2 and TGC3 defined in Figure 2.13 cover the range of $1.05 < |\eta| < 2.4$, while TGC1 covers the range of $1.05 < |\eta| < 2.7$.

Muon trigger chambers

Resistive Plate Chamber (RPC)

The muon trigger detector in the barrel region consists of three stations of RPC; from inside to outside RPC1, RPC2, and RPC3, as shown in Figure 2.11 and 2.13. Each RPC station consists of two active layers measuring η and ϕ. RPC is operated in the avalanche mode, and the signal read-out is performed by strips mounted on the outer faces of the resistive plates. To achieve two-dimensional measurement of (η, ϕ), two sets of strips are assembled orthogonally each other.

Thin Gap Cahmber (TGC)

The TGC is composed of multi-wire proportional chambers (MWPCs) covering the end-cap region ($1.05 < |\eta| < 2.7$, or $1.05 < |\eta| < 2.4$) for triggering. It is composed of four stations, TGC_I, TGC1, TGC2, and TGC3, as shown in Figure 2.13. TGC_I, TGC2, and TGC3 consists of two active layers (Doublet), while TGC1 consists of three active layers (Triplet). Each of the layers performs two-dimensional measurement (η, ϕ) and the redundancy of the track measurements in the layers are used not only to reduce the background contributions or noise hits, but also to mitigate the effects of chamber inefficiency. Thus coincidence with two layers out of three layers and three layers out of four layers coincidence are required in the M1 and M2, M3 stations, respectively. The signal of the TGC_I station is used to require an additional hit to mitigate background events.

Muon precision-tracking chambers

Monitored Drift Tube (MDT)

The precision momentum measurement is performed by the Monitored Drift Tube chambers (MDT). All the MDT chambers consist of two groups of tube layers, called multi-layers, separated by a mechanical spacer. In the innermost station, each multi-layer consists of four tube layers to achieve a better pattern-recognition performance, while in the middle and the outer station, each multi-layer consists of three tube layers. The height of the mechanical spacers depends on the chamber type, varying from 6.5 mm to 317 mm.

Each drift tube is composed of a 50 μm anode wire at the centre of a cathode tube with a diameter of 30 mm. The gas component inside the tube is a mixture of 93 % Ar, 7 % CO_2 and less than 1000 ppm H_2O, where small mixture of water improves high-voltage stability.

The average spatial resolution (z in the barrel region and R in the end-cap region) of individual drift tubes is 80 μm. The position accuracy of the sense wires in a chamber is 20 μm. This results in a chamber spatial resolution of 35 μm.

Cathode Strip Chamber (CSC)

The particle flux is highest in the forward region ($2.0 < |\eta| < 2.7$) thus CSCs, which consists of MWPCs with cathode planes segmented into strips in orthogonal directions, are used in the innermost layer for the range of $2.0 < |\eta| < 2.7$. The CSC chambers have higher rate capability and a timing resolution of 7 ns. Each CSC chamber consists of four wire planes with much finer granularity compared to MDT. The anode wires have a diameter of 30 μm and are made of gold-plated tungsten with 3% rhenium. The position resolution of CSC is 65 μm at 1 kHz/cm^2 in the bending plane and about 5 mm in the transverse plane.

2.2.4 Trigger and data acquisition system

At the interaction point, proton bunches collide at a rate of 40 MHz during stable runs of LHC. The trigger and data acquisition (TDAQ) system takes an important role to decide whether to record an collision event for further processes. Figure 2.14 shows a schematic view of the TDAQ system. The TDAQ system is composed of two types of triggers: hardware-based first-level trigger (Level-1 trigger) and software-based high-level trigger (HLT).

The Level-1 trigger needs to reduce the rate of events to 100 kHz within a latency of Level-1 Calo/Muon: 2.2/~2.0 μs [66], [67]. To achieve this high-speed event processing, the Level-1 trigger exploits the coarse granularity information from electromagnetic and the hadronic calorimeters (Level-1 Calo) as well as the dedicated

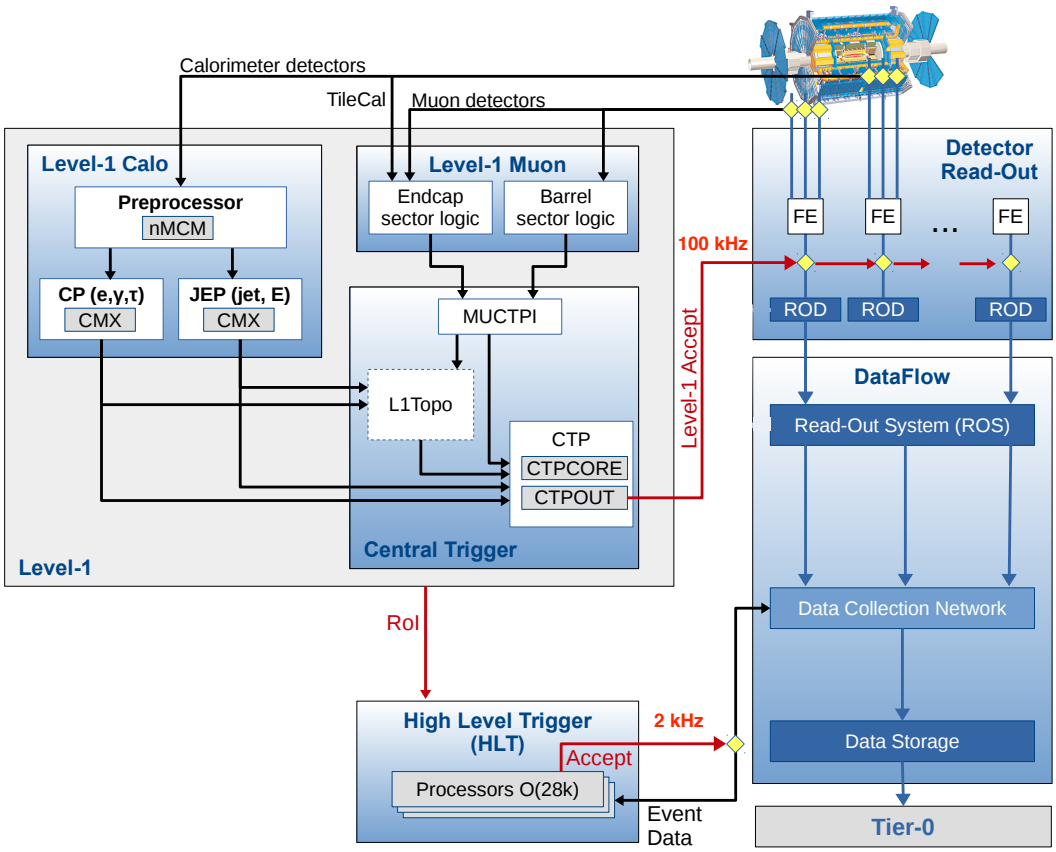


FIGURE 2.14: Schematic view of the TDAQ system composed of a hardware-based Level-1 trigger and a software-based High Level Trigger [65]. L1Topo is being commissioned and not used for the results in this thesis.

muon trigger chambers (Level-1 Muon: RPC and TGC). The level-1 trigger decision is formed by the Central Trigger Processor (CTP) combining the Level-1 Calo and Level-1 Muon information. In the Level-1 Calo trigger, electrons, photons, jets and tau leptons with high transverse-energy (E_T) as well as the high missing E_T candidates are selected. The Level-1 Muon trigger input data from RPC and TGC is fed into the Muon-to-CTP Interface (MUCTPI) [68] to evaluate the total multiplicity of high-p_T level-1 muon candidates. There are six programmable p_T thresholds for the muon candidates and the overlaps across sectors are resolved in the MUCTPI.

The CTP distributes Level-1 Accept (L1A) signal and LHC timing signals to the readout systems of the sub-detector. Upon receiving the L1A signal, (η, ϕ)-coordinates in the detector for each trigger object (the Region-of-Interests, RoIs) are sent to the HLT system. CTP also has a functionality to apply two types of preventive dead-time: 'simple' and 'complex'. The simple dead-time restricts the minimum bunch crossings between two consecutive L1A to avoid overlapping of the readout windows. The complex dead-time adopts a leaky bucket model to avoid the front-end buffers from overflowing; the number of L1A is restricted to 'X' in a given number of 'L' bunch-crossings.

The HLT system defines a final on-line event-selection and is allowed to reconstruct physics objects around RoIs or the event for the full detector. By the regional or full event-reconstruction, more precise trigger selection can be formed in HLT than in the Level-1 trigger. After events are accepted in HLT, all the trigger and the detector information is dumped into a local storage at the experimental site and transferred to the CERN computing centre for off-line reconstruction.

Chapter 3

Data and Monte Carlo simulation

In order to search for the BSM resonances, we compare data collected by the ATLAS detector with Monte Carlo simulation samples including the Standard Model background samples as well as the BSM signals. In this chapter, the dataset and the Monte Carlo simulation samples used for this analysis are presented.

3.1 ATLAS pp collisions data

The full pp collisions data collected in 2015-2016 at the centre-of-mass energy of 13 TeV are used in this analysis, corresponding to an integrated luminosity of 36.1 fb^{-1}. During the 2015 run, LHC increased the number of proton bunches step-by-step, and reached a peak luminosity of 5.0×10^{33} $\text{cm}^{-2}\text{s}^{-1}$, although this peak luminosity is still below the maximum value of LHC Run-1 (7.7×10^{33} $\text{cm}^{-2}\text{s}^{-1}$). In 2016, the peak luminosity exceeded the designed value of 1.0×10^{34} $\text{cm}^{-2}\text{s}^{-1}$ by early summer, and reached up to 1.4×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. The integrated luminosity recorded in the 2015 (2016) pp runs is 3.9 (35.6) fb^{-1}, with an average number of interactions per bunch crossing of about 13.4 (25.1).

The data taking efficiency, which is calculated as the ratio between the recorded luminosity and the total delivered luminosity, is 92.9 % and 92.5 % in the 2015 and the 2016 run, respectively. In order to guarantee the data quality (DQ), it is required that all the sub-detectors are working stably [69] and the LHC beams are stable during the data-taking period. The DQ efficiency, which is defined as luminosity weighted relative detector up-time and good data quality efficiencies, for the 2015 and the 2016 datasets is shown in Table 3.1. The individual efficiency for each of the subsystems is generally better than 99 %, with the exception of the toroid magnet system, which experienced multiple failures, and the Pixel detector in the 2015 dataset, where the IBL was turned off for two runs. In consequence, the total data-taking efficiency for the 2015 and the 2016 datasets is 87.1 % and 95 %, respectively.

3.1.1 Luminosity measurement

The time-integrated luminosity, which represents the amount of data, is calculated from the instantaneous luminosity L. The instantaneous luminosity is calculated from a bunch luminosity L_b which is evaluated by a measurement

TABLE 3.1: Data quality efficiency for the ATLAS pp collision runs in 2015 and 2016 [70]. For each of the detector components, luminosity weighted relative detector up-time and good data quality efficiencies are provided. The efficiencies are calculated for the pp collision runs during the stable beam with 25 ns bunch spacing at $\sqrt{s}$ = 13TeV. As the 2015 dataset, we use the runs between August and November 2015 corresponding to an integrated luminosity of 3.7 fb^{-1}; as the 2016 dataset, we use the runs between April and October 2016 corresponding to an integrated luminosity of 35.9 fb^{-1}. The total efficiency is calculated requiring that all the detector components work without any errors. For the 2015 run, the lower DQ efficiency in the Pixel detector is due to the IBL being turned off for two runs. In the 2016 run, the toroid magnet was off for some runs, leading a loss of the efficiency.

Year	Inner Detectors			Calorimeters		Muon Spectrometer				Magnets		Trigger	Total
	Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid	L1	
2015	93.5	99.4	98.3	99.4	100	100	100	100	100	100	97.8		87.1
2016	98.9	99.9	99.7	99.3	98.9	99.8	99.8	99.9	99.9	99.9	97.2	98.3	93-95

TABLE 3.2: Summary of Monte Carlo simulation samples both for the BSM signals and the SM backgrounds.

Signal/Background	Sample	Generator	PDF
Signal	Heavy neutral Higgs boson	POWHEG+PYTHIA	CT10 + CTEQ6L1
	Heavy Vector Triplets	MADGRAPH+PYTHIA	NNPDF23_lo
	KK graviton	MADGRAPH+PYTHIA	NNPDF23_lo
Background	$W \to \ell\nu$+jets	SHERPA	NNPDF30_nnlo
	$W \to \ell\nu$+jets, alternative	MADGRAPH+PYTHIA	NNPDF23_lo
	$Z \to \ell\ell$+jets	SHERPA	NNPDF30_nnlo
	$Z \to \ell\ell$+jets, alternative	MADGRAPH+PYTHIA	NNPDF23_lo
	$t\bar{t}$	POWHEG+PYTHIA	CT10
	$t\bar{t}$, alternative 1	POWHEG+HERWIG++	CT10
	$t\bar{t}$, alternative 2	MADGRAPH	CT10
	Single top	POWHEG+PYTHIA	CT10
	Dibosons	POWHEG+PYTHIA	CT10NLO+CTEQ6L1

of the cross section (σ_{inel}) and event rate of the inelastic scattering (R_{inel}) of the pp collision:

$$L_b = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{\mu f_{\text{rev}}}{\sigma_{\text{inel}}} \quad , \tag{3.1}$$

where μ is the number of inelastic interactions in the bunch crossing, and f_{rev} is the bunch revolution frequency. The total instantaneous luminosity is given taking the sum of L_b over the number of bunches colliding at the interaction point n_b [71]:

$$L = n_b \langle L_b \rangle = n_b \frac{\langle \mu \rangle f_{\text{rev}}}{\sigma_{\text{inel}}} \quad , \tag{3.2}$$

where $\langle L_b \rangle$ is the mean bunch luminosity and $\langle \mu \rangle$ is the bunch-averaged pile-up parameter, which is measured with a dedicated detector at the forward region. n_b and f_{rev} are given from the LHC operational information. The measured pile-up parameters and the measured cross section are corrected with the efficiency of the detector. The luminosity measurement is calibrated to the inelastic cross section (assumed to be $\sigma_{\text{inel}} = 80$ mb for 13 TeV [59]) using beam-separation scans, also known as the van der Meer method [72]. In this method, the effective transverse beam size is measured by performing scans in separation. As a result, we obtain the curve of the interaction rate versus the separation. The uncertainty on the integrated luminosity for the 2015+2016 dataset is 3.2%.

3.2 Monte Carlo simulation

The simulated data of signals and backgrounds are used to optimise the event selection and to estimate the SM background contributions. The Monte Carlo (MC) simulation samples used in this analysis are summarised in Table. 3.2. In order to reproduce the distribution of the number of interactions per bunch crossing in the data, the MC samples are reweighted by overlaying additional pp collisions generated with PYTHIA 8.186 [73] on the samples. The accuracy of the average correction factor due to pileup reweighting is found to be ±3%. The produced samples are processed through the full ATLAS detector simulation based on GEANT4, which is explained in Subsection 3.2.5.

3.2.1 Parton Distribution Function

In LHC, the inelastic pp collision is regarded as point-like interaction between partons. The parton distribution function (PDF) plays an important role to calculate the cross section of the pp collision as it gives the probability of finding a parton in the proton with flavour i (i denotes quarks or gluon) carrying a fraction x of the proton momentum Q ($p = xQ$, where p is the parton momentum, here x is called "Bjorken scale"). Cross section of a hard process of any hadronic final states in the pp collisions $\sigma_{pp\to X}$ is factorised into the hard part ($\hat{\sigma}_{pp\to X}(x_1, x_2, \mu^2, Q^2)$) and

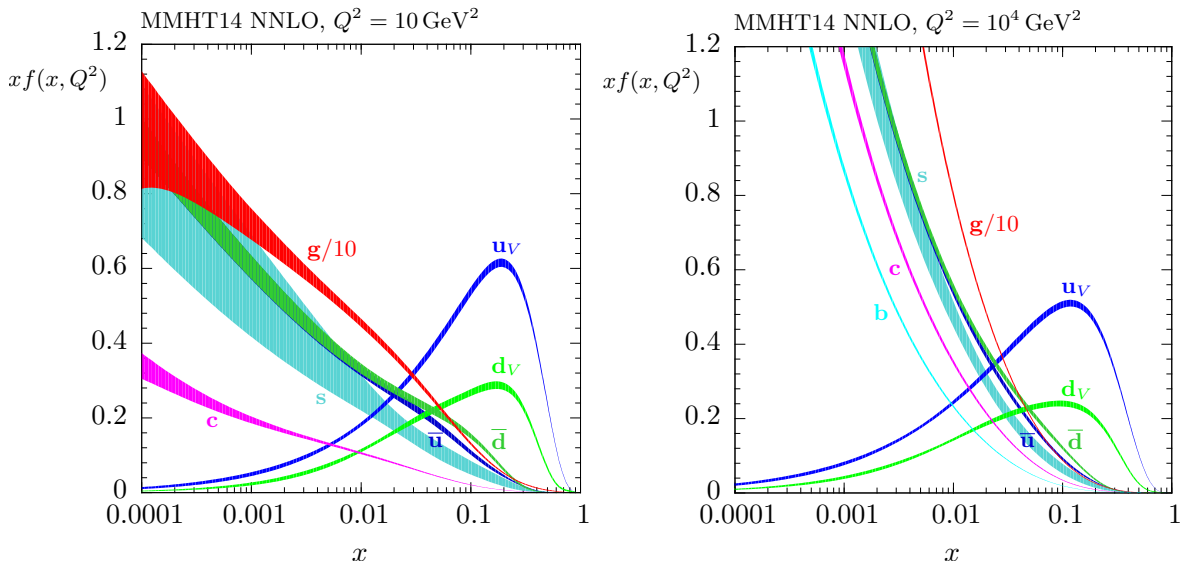


FIGURE 3.1: MMHT2014 NNLO PDFs at $Q^2 = 10\ \mathrm{GeV}^2$ (left) and $Q^2 = 10^4\ \mathrm{GeV}^2$ (right) with associated 68 % confidence-level uncertainty bands [75]. The blue and green lines show the distributions for u and d quarks, which make peaks at the high x region. In the low x region, the contributions from gluon and the sea quarks are dominant. The gluon distribution is shown with a scale factor of 1/10.

normalisation from PDFs:

$$\sigma_{pp\to X} = \sum_{i,j} \int dx_1 dx_2 f_i(x_1,\mu^2) f_j(x_2,\mu^2) \times \hat{\sigma}_{pp\to X}(x_1,x_2,\mu^2,Q^2) \quad , \tag{3.3}$$

where f_i is a PDF (i denotes its flavour). The sum is calculated over all contributing parton sub-processes $pp \to X$ contained in the perturbatively calculable cross sections in QCD $\hat{\sigma}_{pp\to X}$ [74]. The factorisation scale, which is thought as a scale that separates the long- and short-distance physics, and the renormalization scale for the running QCD coupling is usually assumed to be a same value (μ) at the momentum scale of the hard-scattering process.

PDFs can not (yet) be predicted by first principles with non-perturbative QCD calculation. The perturbative QCD, however, accurately predicts the evolution of PDFs in Q^2. Thus, it is necessary to use experimental measurements to determine probability of the initial state: the PDFs $f(x, Q_0^2)$ at a given starting scale of the order $Q_0^2 \sim 1$ GeV^2.

Many fixed target and collision experiments provided data for PDF and global fit results to the data. Figure 3.1 shows the next-to-next-to-leading order (NNLO) perturbative QCD calculation of the $xf(x, Q^2)$ by the MMHT collaboration [75] as a function of x at two benchmark Q^2 values. There are peaks at the higher x region for the valence quarks (u, d), whereas the gluons and the sea quarks are dominant at the small x region. This feature is remarkably shown at the high Q^2 region, which is the main phase space at LHC ((100GeV)^2). In the following sections, we explain the PDF sets for the MC samples used in the analysis.

3.2.2 Event generators

We use multiple MC generators in this thesis. In this subsection, the generators used in this thesis are summarised with a brief description.

PYTHIA

Pythia is a general-purpose generator. This is one of the most common generators for physics analyses at LHC. The version 6 of Pythia [76] is written based on Fortran 77 while the version 8 [73] is completely rewritten by C++. In this thesis, we use both versions; the version 8 provides a modern description of the physics, the version

6 is well tested. Pythia generates hard processes at the leading-order (LO)[1] up to the parton shower, hadronisation, underlying event calculation, and particle decays. In this thesis, we interface Pythia to the other event generators to calculate hard processes at higher orders.

Powheg

Powheg [77] calculate hard processes at the next-to-leading-order (NLO) while it is not able to simulate the whole chain from calculating hard processes to the particle decays. In this thesis, this generator is interfaced with Pythia and Herwig++ to perform the modelling of the parton showers, hadronisation, modelling of the underlying event, and particle decays.

Sherpa

Sherpa [78] is a general-purpose generator, capable of simulating the processes in lepton-lepton, lepton-hadron, and hadron-hadron collisions as well as photon induced processes. This generator is also possible to calculate the matrix elements at NLO.

Herwig++

Herwig++ [79] is also a general-purpose generator. This generator calculates hard processes at LO and simulates decays with full spin correlations. This generator also provides a precise description of hadronic decays, particularly for bottom hadrons and τ leptons.

MadGraph

MadGraph [80] automatically generates the amplitudes for all the relevant sub-processes, and provides a stand-alone code to calculate cross sections. We use Pythia as a alternative sample to calculate the background modelling uncertainties. This generator is also used interfaced to Herwig++ and Pythia to perform the modelling of the parton showers, hadronisation, modelling of the underlying event, and particle decays.

3.2.3 Standard Model backgrounds

The dominant SM background for the di-boson resonance signal comes from events with a leptonically decaying W boson produced in association with jets ($W \to \ell\nu$+jets). The SM background including the production of top quarks ($t\bar{t}$ or single top-quark), multijets, dibosons, and $Z \to \ell\ell$+jets are also considered.

Sherpa v2.2.1 [78] with the NNPDF30_nnlo [81] PDF set is used for the events containing leptonically decaying W or Z bosons with associated jets. In the Sherpa samples, multi-parton matrix elements are calculated for up to 2 partons at NLO and 4 partons at LO using a matrix element generator, namely Comix [82], and OpenLoops [83], by which we calculate scattering amplitudes at one loop, and merged with the parton shower in Sherpa for bulk production [84] using the ME+PS@NLO prescription [85]. The W/Z+jets samples are sliced to several ranges with max $\{h_{\mathrm{T}}, p_{\mathrm{T}}(V)\}$ and the number of b and c quarks in the final state; here, h_{T} is the scalar sum of p_{T} of jets in the event and $p_{\mathrm{T}}(V)$ is the transverse momentum of the W/Z boson. To estimate systematic uncertainties related to the V+jets processes, alternative samples are produced using MadGraph5_aMC@NLO 2.2.2 [80] with an interface to the Pythia 8.186 parton shower model, and the A14 tune together with the NNPDF23_lo PDF [86].

For the generation of $t\bar{t}$ and single top-quarks, the Powheg-Box v2 generator [77] with the CT10 PDF sets in the matrix element calculations is used. In all these processes, top-quark spin correlations are preserved, and the top mass is set to be 172.5 GeV. The parton shower, fragmentation and the underlying event are simulated using Pythia 6.428 [76] with the CTEQ6L1 [87] PDF sets and the corresponding Perugia 2012 tune (P2012) [88]. The renormalisation and factorisation scales are set as: $\mu = 4(m_b^2 + p_{\mathrm{T}}(b)^2)^{1/2}$ for t-channel; where b denotes the spectator b-quark, $\mu = m_t$ for s-channel and Wt channel, and $\mu = (m_t^2 + p_{\mathrm{T}}(t)^2)^{1/2}$ for $t\bar{t}$. Systematic uncertainties associated with parton showering and hadronization are evaluated using alternative Powheg-Box samples with an interface to Herwig++ 2.7.1 [79] using the UEEE5 underlying event tune [89]. A background modelling uncertainty is computed using different samples produced with MadGraph5_aMC@NLO 2.2.2 [80] with the CT10 PDF.

The diboson productions (WW, WZ, and ZZ) with one of the bosons decaying hadronically and the other leptonically are simulated using the Sherpa 2.2.1 generator with the CT10 PDF set. Additional diboson events using the Powheg-Box generator, with an interface to the Pythia 8.186 parton shower model, are generated for estimating

[1] Leading-order(LO): tree level, Next-to-leading-order (NLO): up to one-loop level, and Next-to-next-to-leading-order (NNLO): up to two-loop level

systematic uncertainties. The CT10NLO set is used for the PDF of the hard-scatter process and the CTEQ6L1 PDF is used for the parton shower.

The cross sections of all the background processes are computed at the next-to-next-leading order (NNLO) in QCD [90–94], except for the diboson samples for which the generator cross sections at NLO are used. The EvtGen v1.2.0 program [95] is used for simulating the bottom- and charm-hadron decays, except for samples generated by Sherpa for which multi-parton matrix elements are calculated using Comix and OpenLoops.

3.2.4 Signals

Signal samples for the heavy neutral Higgs boson are generated for spin-0 interpretation with the Narrow Width Approximation (NWA) using Powheg-Box v1 [77, 96] with NLO gg_H [97] and VBF_H [98] modules and the CT10 PDF set [99]. The Powheg-Box event generator is interfaced to Pythia 8.186 [73] for parton showering, hadronisation, and underlying event using the AZNLO set [100] and the CTEQ6L1 PDF [87]. Possible effects from an interference between the SM diboson backgrounds and the low-mass Higgs boson at m_H =125 GeV are negligible. This benchmark is used for a feasibility study of searches for extended Higgs sectors. WW resonance signals with masses ranging from 300 GeV to 3000 GeV are produced via either the gluon-gluon fusion and vector-boson fusion processes separately [101, 102].

The Heavy Vector Triplet (HVT) benchmark signals, $Z' \to WW$ and $W' \to WZ$ [28, 29], are generated mainly via $q\bar{q}$ annihilation for spin-1 interpretation. The HVT signals introduce a phenomenological Lagrangian in which the couplings of a triplet vector boson with the SM fields are parametrized. The coupling between the new triplet and the SM fermions is given by the combination of parameters $g^2 c_F / g_V$, where g is the SM $SU(2)_L$ gauge coupling, c_F is a multiplicative factor that modifies the coupling to fermions, and g_V represents the coupling strength of the known W and Z bosons to the new vector bosons. The coupling between the Higgs boson and the new triplet is given by the combination $g_V c_H$ where c_H is a multiplicative factor that modifies the coupling to the Higgs boson. Two different benchmark signal models are considered in this thesis, namely 'model-A' and 'model-B' corresponding to $g_V = 1$ and $g_V = 3$, respectively. The HVT model-A is typical as an extended gauge model [103] with the heavy vectors having comparable branching ratios into fermions and gauge bosons. The HVT model-B is representative of composite Higgs models, where the fermionic couplings are suppressed [104–106]. In both scenarios, the width of the decay is narrower than the detector resolution. The VBF production can be enhanced if the coupling to the fermions is suppressed for some reason, thus the HVT VBF samples are generated with the coupling to fermions set to be zero, and the couplings to the gauge bosons similar to those of the HVT model-A. The signal samples are produced in a mass range from 300 GeV to 5000 GeV using MadGraph5_aMC@NLO 2.2.2 [80] interfaced to Pythia 8.186 with the NNPDF23_lo [86] PDF and the A14 tune [107] for the underlying event.

The KK graviton decaying to WW is produced via the ggF process with the mass range from 300 GeV to 5000 GeV using MadGraph5_aMC@NLO with a model given in Reference [108]. G_{KK} is the first Kaluza–Klein mode [109] of a spin-2 graviton in a warped extra dimension with a curvature parameter k and a dimensionless coupling constant $k/\bar{M}_{Pl} \sim O(1)$, where $\bar{M}_{Pl}$ is the effective four-dimensional Planck scale ($\bar{M}_{Pl} = 2.4 \times 10^{18}$ GeV). We interpret results with two bulk RS G_{KK}: with $\bar{M}_{Pl} = 1.0$, and with $\bar{M}_{Pl} = 0.5$, to compare from the ones of the CMS collaboration. The signal samples with $\bar{M}_{Pl} = 0.5$ are reproduced from the signal samples with $\bar{M}_{Pl} = 1.0$ taking into account the different widths between $\bar{M}_{Pl} = 1.0$ and $\bar{M}_{Pl} = 0.5$. However, the integrated luminosity of the dataset is not large enough to evaluate an upper-limit of the cross-section for this resonance. Table. 3.3 summarises the width of the resonances and the cross-sections and branching ratios for HVT and the KK graviton.

3.2.5 Detector simulation

After the event generation, each particle is propagated through the full ATLAS detector simulation by Geant4 [110, 111]. Geant4 provides packages to calculate the trajectory of particles in the detector geometry. The simulation results are compressed to store what is necessary for physics analyses and detector-performance studies. The same online trigger algorithms used for the data acquisition in ATLAS runs on the simulation samples. After the digitisation of the hits in the detectors according to the detector responses, the output of the simulation is finally transferred into the identical format with the data recorded in ATLAS. Thus, both the MC and real data samples can be analysed with same software packages.

TABLE 3.3: The resonance width (Γ) and the cross-section times branching ratio ($\sigma \times BR$) for diboson final states, for different values of the pole mass m for a representative benchmark for spin-1 and spin-2 cases. The table shows the predictions by model-B of the HVT parameterization ($g_V = 3$) and by the KK graviton model ($k/\bar{M}_{\mathrm{Pl}} = 1$).

	HVT W′ and Z′			G_{KK}	
m	Γ	$\sigma \times BR(Z' \to WW)$	$\sigma \times BR(W' \to WZ)$	Γ	$\sigma \times BR(G_{\mathrm{KK}} \to WW)$
[TeV]	[GeV]	[fb]	[fb]	[GeV]	[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

Chapter 4

Event reconstruction and particle identification

In this chapter, we summarise the procedures and methodologies employed for the reconstruction and identification of the physical objects; vertex, track, electron, photon, muon, jet, and the missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$). For each of the objects, performance for momentum/energy measurements will be provided.

4.1 Track and vertex reconstruction

The most general requirement on an signal event is to have a 'primary vertex', defined as a vertex with two or more associated charged-particle trajectories (tracks) carrying the largest sum of the squared transverse momentum of all the tracks associated to the vertex. Tracks are reconstructed with the hits in the inner detector and used in the electron, muon, and jet reconstruction algorithms. The kinematic variables used in analyses are calculated with respect to the primary vertex, thus efficient reconstruction of the primary vertices is essential for determining the full kinematic properties of a hard-scattering interaction and of soft interactions as a measure of amount of the pile-up.

4.1.1 Track reconstruction

The ATLAS track reconstruction algorithm [112] begins with assembling clusters from the raw Pixel and SCT hits where the deposited energy yields a charge above threshold. Figure 4.1 shows a schematic view of the track reconstruction. We create three-dimensional measurements referred to as 'space-points', which represent the points where the tracks traversed in the active material of the inner detector from these clusters, then we reconstruct a 'track seed' using all the possible combinations of three space-points. The quality of tracks significantly depends on the purity, or fraction of seeds, which varies depending on which sub-detectors record the space-points used in the seed. Thus, we consider 'seed types' categorised into three; SCT-only seeds, pixel-only, and mixed-detector seeds, from higher quality to the lower quality, representing the order of purity. Quality is further improved by associating one additional space-point that is compatible with the trajectory. We use a combinatorial Kalman filter [113] to reconstruct track candidates from the track seeds by associating additional space-points from the remaining layers of the pixel and SCT detectors. If there are more than one compatible space-points on the same layer, the filter creates multiple track candidates.

We calculate a 'score' for each of the track candidates, then tracks with bad score are rejected from the candidates. The score is increased when clusters are assigned to a track and is decreased when the track intersects a detector hole. The track candidates are rejected if they fail to meet any of the following basic quality criterion:

- $p_{\mathrm{T}} > 400$ MeV,
- $|\eta| < 2.5$,
- At least of 7 pixel and SCT clusters (12 are expected),
- One shared pixel cluster or at most two shared SCT clusters on the same layer,
- Not more than two holes in the pixel and SCT detectors,
- Not more than one hole in the pixel detector, here a hole is defined as a missing measurement when it is expected given the predicted tracks,
- d_0^{BL}[1]< 2.0 mm,
- $|z_0^{\mathrm{BL}}$[2]$\sin\theta < 3.0|$ mm,

[1]The transverse impact parameter calculated with respect to the measured beam-line position. Schematically illustrated in Figure 4.5(a)

[2]The longitudinal difference along the beam line between the point where d_0^{BL} is measured and the primary vertex, and θ is the polar angle of the track. Schematically illustrated in Figure 4.5(b)

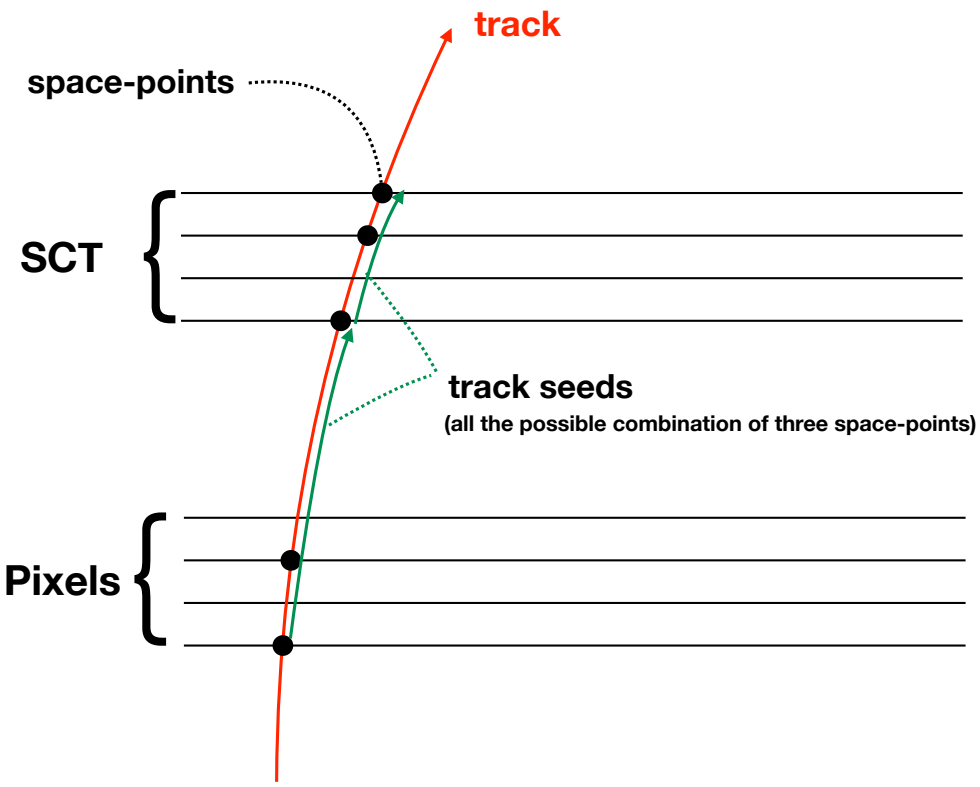


FIGURE 4.1: Schematic view of the track reconstruction. Definitions of space-points and track seeds are described in the main body of this section.

For the track candidates which pass the quality tests, we perform a fit using all the available track information, then the tracks in the result of the fitting are added to the final track collection. We perform this track fit in the final stage, thereby we minimise the number of times to call the fitter.

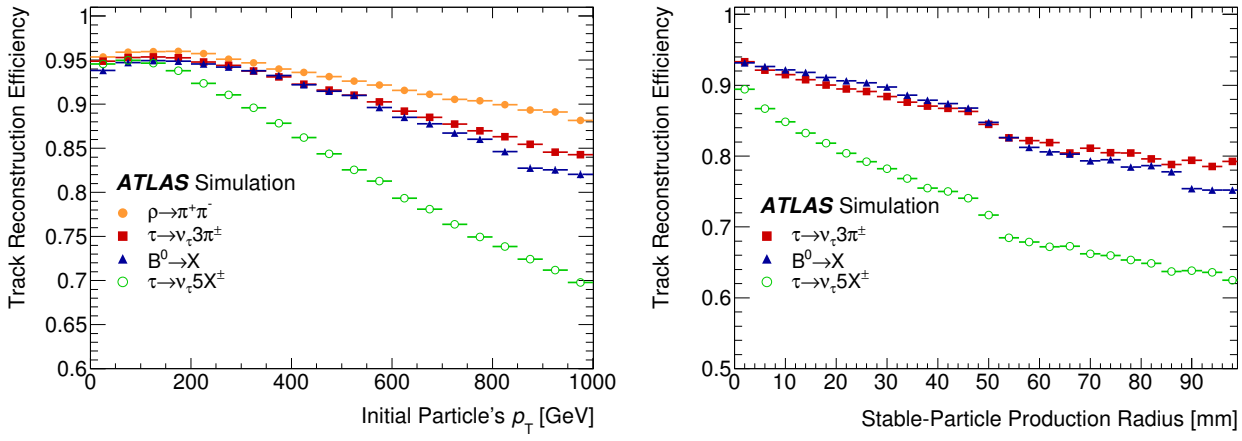
The plots in Figure 4.2 show the track reconstruction efficiency as a function of particle p_{T} and production radius. The production radius is defined as the radial distance between the decay vertex of the parent particle and the beam axis. A small inefficiency at the low initial particle's p_{T} region is due to inelastic interactions, such as hadronic interactions. The efficiency decreases due to the increasingly collimated decay products at high initial particle's p_{T}. When the separation between highly collimated charged particles is comparable to the granularity of the inner sector, we can fail to reconstruct the tracks. At the higher production radius region, the efficiency also decreases because the charged particles arrive at each active layer with less separation. It is important to optimise the track quality cuts for each signature (electrons, muons, or hadrons) with the signal efficiency and fake probability. The track quality conditions for electrons and muons are explained in the following sections.

4.1.2 Vertex reconstruction

We reconstruct the vertices using the reconstructed tracks. The tracks must satisfy the following selections to be used in the vertex reconstruction [114]:

- $p_{\mathrm{T}} > 400$ MeV and $|\eta| < 2.5$,
- Number of hits in Pixel and SCT $\geq 9(11)$, $|\eta| < 1.65$ ($|\eta| > 1.65$) .
- At least one hit in the first two pixel layers,
- A maximum of 1 shared module (1 shared pixel hit or 2 shared SCT hits),
- No detector hole in Pixel, less than two holes in SCT.

Using the tracks which satisfy the selection criteria, vertices are reconstructed in the following steps, which are also illustrated in Figure 4.3 [114]:



(a) (b)

FIGURE 4.2: Single-track reconstruction efficiency as (a) a function of the initial particle's p_{T} for ρ, three- and five-prong τ and a B^0 which decays before the IBL, and (b) versus the production radius (defined in the main body of this section) for the decay products of a three- and five-prong τ as well as a B^0, where no requirement imposed on the production radius of stable charged particles. [112]

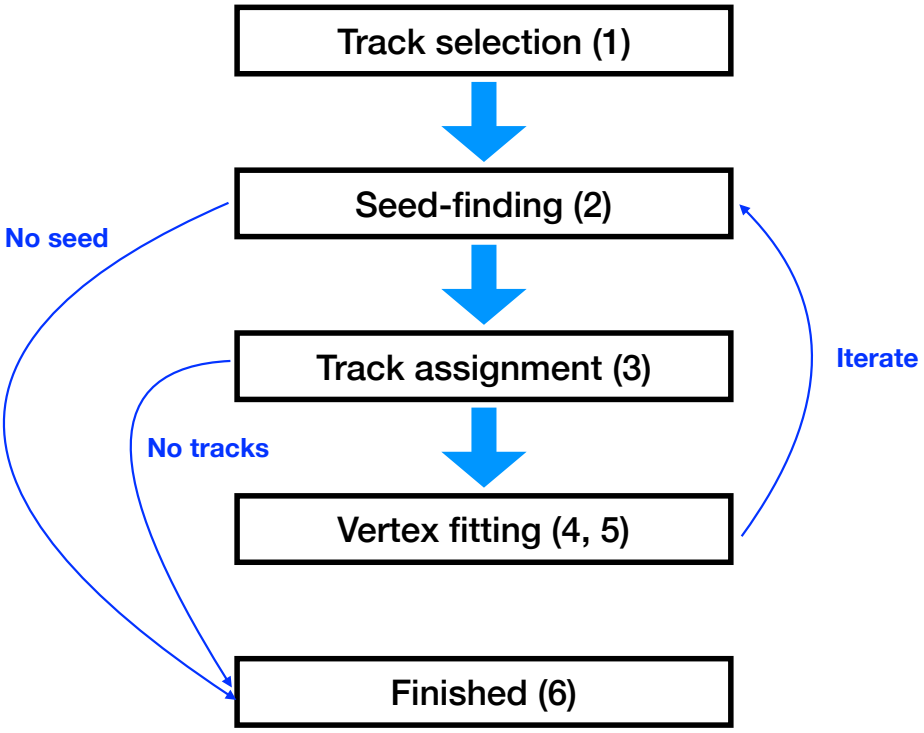


FIGURE 4.3: Vertex reconstruction flow. The number in a pair of parentheses represents the step number.

1. A set of tracks passing the vertex selection criteria is collected.
2. The first vertex position is selected as a 'seed position'. The transverse position of the seed is taken at the centre of the beam line; the vertical position is estimated from tracks, considering the track points of closest approach to the beam spot.
3. Tracks compatible with the seed are collected for fitting.
4. The tracks and the seed are used to fit the best vertex position; we perform a χ^2 minimisation fit iteratively to find the optimal vertex position, using the seed position as the initial point and kinematics of reconstructed tracks as input parameters. We assign a weight to each track reflecting its compatibility with the vertex estimate.
5. After the vertex position is determined, tracks that are incompatible with the vertex are removed from the correspondence list and allowed to be used to determine another vertex. The vertices are required to have at least two associated tracks.
6. The procedure 2-5 is repeated until no unassociated tracks are left in the event or no additional vertex can be found in the remaining tracks.

The output of the vertex reconstruction is a set of three dimensional vertex positions.The vertex with the largest scalar sum of associated track p_{T} is chosen as the primary vertex for the hard scattering.

The vertex reconstruction efficiency is measured in data as a function of number of tracks associated to a vertex using a low-pileup ($\mu < 0.01$) dataset. The efficiency is evaluated by taking the ratio of the number of events between the cases with a reconstructed vertex as numerator and with at least two reconstructed tracks as denominator. The reconstruction algorithm becomes fully efficient for the events with five or more tracks as shown in Figure 4.4(a).

The number of reconstructed vertices scales linearly with μ in the ideal condition of perfect reconstruction efficiency. However, there are several cause of non-linearity as shown in Figure 4.4(b). The most significant effect is vertex merging; two or more interactions are reconstructed as one vertex. Secondary effects come from fake and split vertices, reconstruction inefficiencies, detector acceptance, and non-collision backgrounds. Taking into account the non-linear effects, an empirical model to compute the average number of reconstructed vertices, $\langle n_{\mathrm{Vertices}} \rangle$ as a function of μ is determined as follows [114]:

$$\langle n_{\mathrm{vertices}} \rangle = p_0 + \epsilon \cdot \mu - F(\epsilon \cdot \mu, p_{\mathrm{merge}}) \quad , \tag{4.1}$$

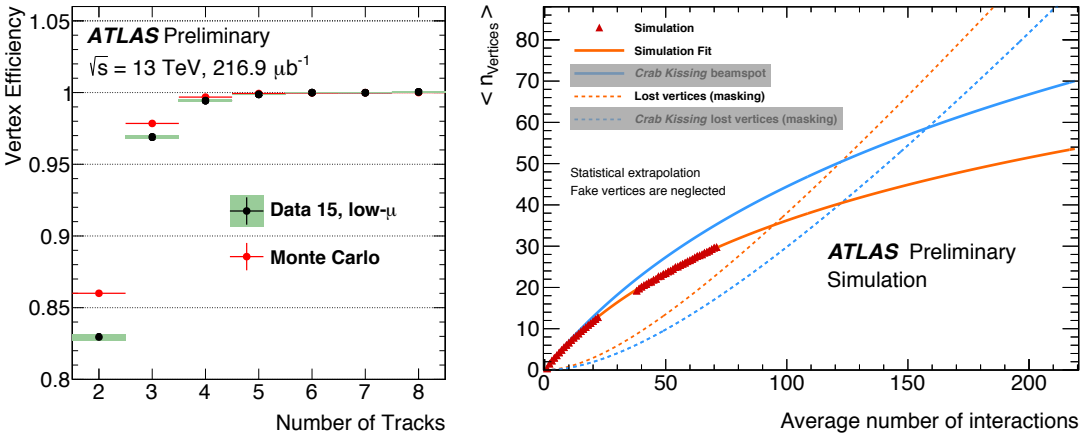
where ϵ is the efficiency of the vertex reconstruction algorithm, and p_0 accounts for any small offset arising from non-collision backgrounds. The quantity $\epsilon \cdot \mu$ represents the linearity to the average number of vertices that would be reconstructed in the absence of any pile-up induced vertex merging effects. The function $F(\epsilon \cdot \mu, p_{\mathrm{merge}})$ represents the average number of vertices lost by the merging effects, taking into account the vertex merging probability, p_{merge}, derived from a low-μ data sample ($\mu < 0.01$, used in Figure 4.4(a)), and is used to derive a final prediction, shown in Figure 4.4(b).

4.1.3 Impact parameter resolution

The primary vertex is used as a reference point to evaluate the transverse and longitudinal impact parameters, as shown in Figure 4.5(a) and 4.5(b), respectively. The resolution of the impact parameters indicates the performance of the track and vertex reconstruction. Figure 4.5(c) and 4.5(d) show d_0 and $z_0 \sin\theta$ distributions, respectively. We observe small discrepancies between data and MC because of two reasons; (1) material description used in simulation for the IBL, (2) the residual misalignment of the detector components. The resolution is defined as the width of a Gaussian fit and it depends on p_{T} and η.

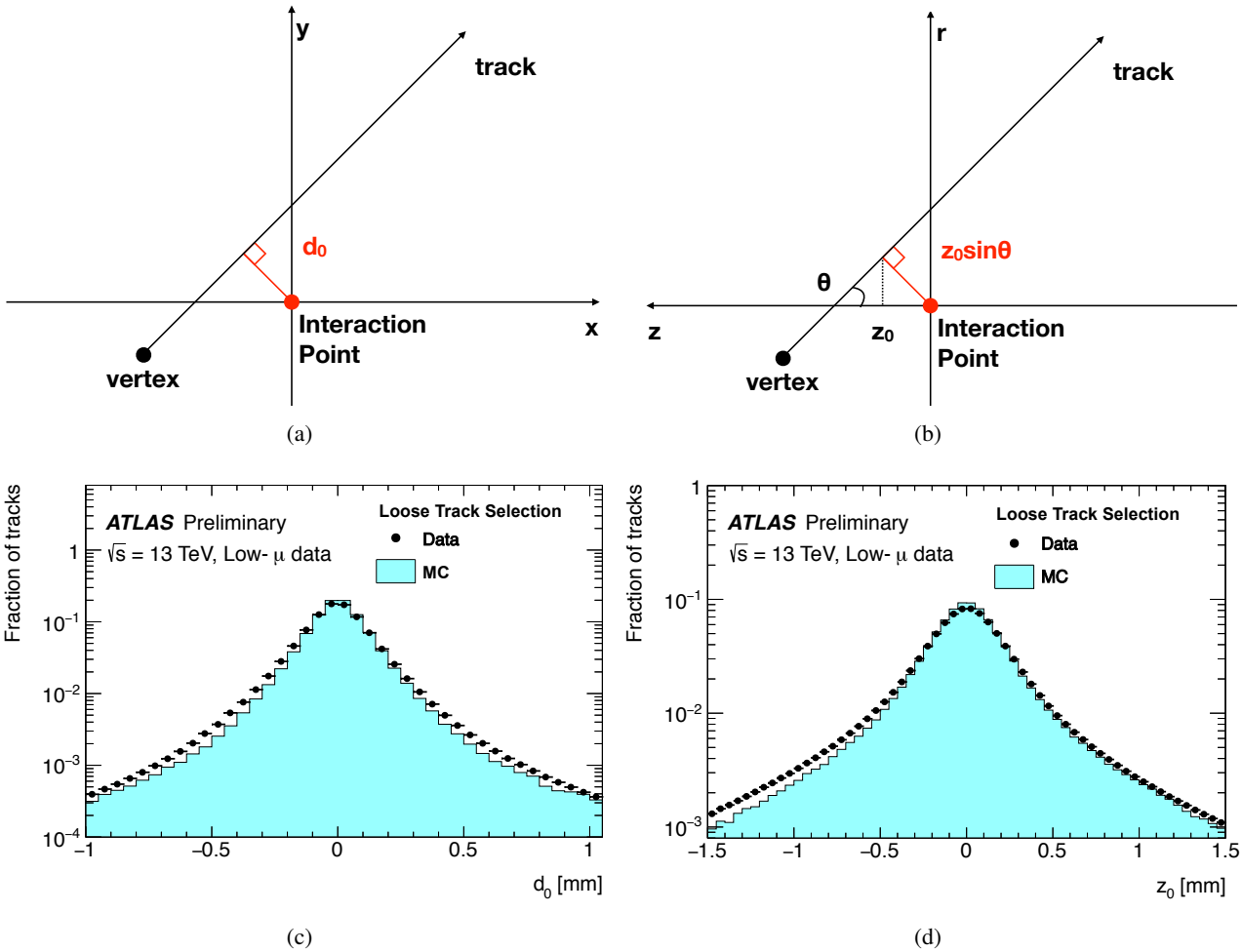
4.2 Electron reconstruction

The electron reconstruction starts with the information provided by the electromagnetic (EM) calorimeter. The information from the calorimeter is combined with the tracks in the inner detector; electrons are identified using a likelihood. Energy calibration is applied based on the MC generator information ("truth") and validated by using the data of $Z \to ee$ events.



(a) (b)

FIGURE 4.4: (a) Efficiency for vertex reconstruction as a function of the number of tracks in a low-μ ($\mu < 0.01$) dataset. The shaded error bars represent the statistical uncertainty. The mild differences between data and prediction are in MC due to the MC tuning, which was optimised for higher track multiplicities. (b) The average number of reconstructed vertices as a function of the average number of interactions μ. The red triangles and the orange line show the prediction in MC for the parameterisation. The blue lines are not used in this thesis. The number of lost reconstructed vertices by masking effects is shown by the dotted line [114].



(a) (b)

(c) (d)

FIGURE 4.5: Definitions for (a) transverse and (b) longitudinal impact parameters, indicated by red lines. Distributions of (c) transverse (d_0) and (d) longitudinal (z_0) impact parameters, with respect to the primary vertex. These results are derived using a low-μ dataset ($\mu < 0.005$) [115].

4.2.1 Seed-cluster reconstruction

For the first step, we form a 'seed-cluster' in a 'sliding window' algorithm [116] to find energy deposits by electrons, as following. The EM calorimeter is divided into an η-ϕ matrix ('cells') with $N_\eta = 200$ and $N_\phi = 256$. The sliding window has a fixed-size of cells, 3×5 in units of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ corresponding to the granularity of the middle EM calorimeter. The energy in the cells is summed to search for electron cluster seed with total transverse energy of the cluster larger than 2.5 GeV. Then, we form the 'seed-clusters' around the reconstructed seeds using a clustering algorithm [116].

4.2.2 Track matching for electrons

We search for tracks that are compatible with the reconstructed seed-clusters. Tracks are extrapolated from the ID to the middle EM calorimeter, then we apply the following selections on $\Delta\eta$ and $\Delta\phi$, which are the differences of the azimuthal angle and pseudo rapidity between the ID and the EM calorimeter; $\Delta\phi < 0.1$ toward the track-bending direction or $\Delta\phi < 0.05$ in the opposite direction; $\Delta\eta < 0.05$ at the seed-cluster. To account for energy losses due to bremsstrahlung, this asymmetric $\Delta\phi$ selection is used. If multiple tracks match a seed-cluster, they are sorted in the following order, with the first being used to define the electron properties.; tracks with hits in the pixel detector are preferred, then tracks with hits in the SCT but not in the pixel detector. Within each category, tracks with smaller $\Delta R (= \sqrt{\Delta\eta^2 + \Delta\phi^2})$ are preferred. Tracks associated to the seed-clusters are re-fitted using a Gaussian Sum Filter (GSF) [117], which is used to compensate the electron energy losses in the ID due to bremsstrahlung.

4.2.3 Energy calibration

The energy calibration for electrons is divided into two steps; (1) a MC-based calibration and (2) a data-driven calibration. In the first step, we use an approach using a multivariate analysis (MVA) [118] to calibrate the cluster energy to the truth electron energy (E^{MC}) in the MC sample. The MVA is optimised separately for electrons, converted photons, and unconverted photons[3]. The MVA algorithm exploits the following variables as inputs; uncalibrated energy; the ratio between the energies in the pre-sampler and the calorimeter; the ratio between the energies measured in the first two layers of the calorimeter; the cluster barycenter in η; and the cluster barycenter in η and ϕ in the calorimeter frame.

In the next step, we perform further calibration of the electron energy using data to remove residual disagreement between data and simulation [118]. The energy mis-calibration is parametrized as:

$$E_i^{\mathrm{data}} = E_i^{\mathrm{MC}}(1 + \alpha_i) \quad , \tag{4.2}$$

where E_i^{data} and E_i^{MC} are the electron energies in data and simulation, and α_i represents the deviation from optimal calibration, in a given η region labelled i. The relative energy resolution is parametrized as:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad , \tag{4.3}$$

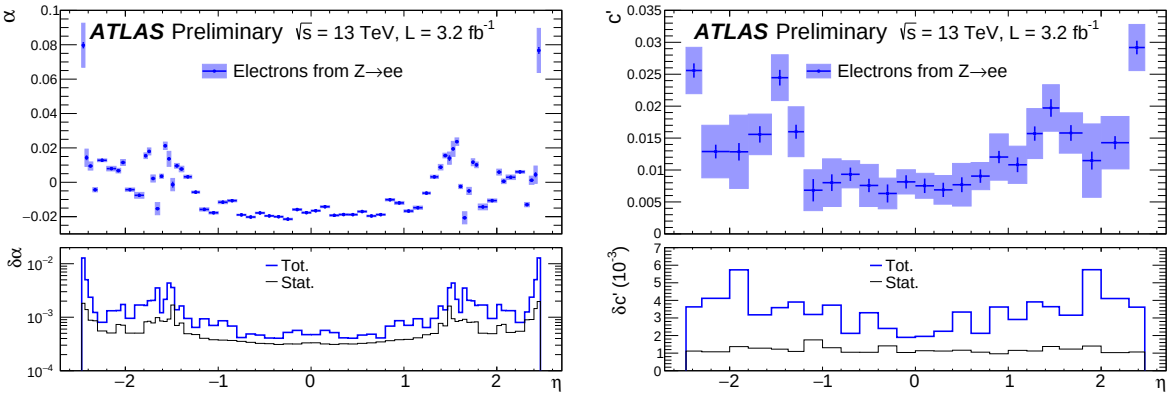
where a is the sampling term related to shower fluctuations in the calorimeter and modelled by simulation, b is the electronic noise term measured in calibration runs, and c is the constant term. The difference in energy resolution between data and simulation is modelled by an additional effective constant term c_1':

$$\left(\frac{\sigma(E)}{E}\right)_i^{\mathrm{data}} = \left(\frac{\sigma(E)}{E}\right)_i^{\mathrm{MC}} \oplus c_i' \quad . \tag{4.4}$$

The sources of systematic uncertainty for the data-driven calibration can be grouped in three categories:

- event selections:
  - election criteria for the electron candidates (identification, isolation and energy loss);
  - differences in the efficiency of electron in the event selection between data and simulation;
  - selection criteria for the candidate events for $Z \to ee$;
  - evaluation of the impact of the backgrounds ($Z \to \tau\tau$ and $t\bar{t}$);

[3] If no tracks are associated to a cluster, it is classified as an unconverted photon. In case a pair of oppositely-charged tracks matches a cluster, it is considered as converted photons.



(a) (b)

FIGURE 4.6: Energy scale factor (a) α and (b) c' for energy resolution in $Z \to ee$ events as a function of η. The uncertainty bands on the top panels show the total uncertainties on these quantities, while the statistical and the total systematic uncertainties are shown separately in the bottom panels [118].

- calibration procedure:
  - methodology (strategies for the extraction of α_i and c'_i);
  - accuracy of the method (the residual bias of this method, estimated using pseudo-data samples is assigned as a systematic uncertainty);
- material:
  - mis-modelling of the material for the IBL in the simulation.

The best fit results for α and c' are shown in Figure 4.6, with the statistical and total systematic uncertainties.

4.2.4 Electron identification

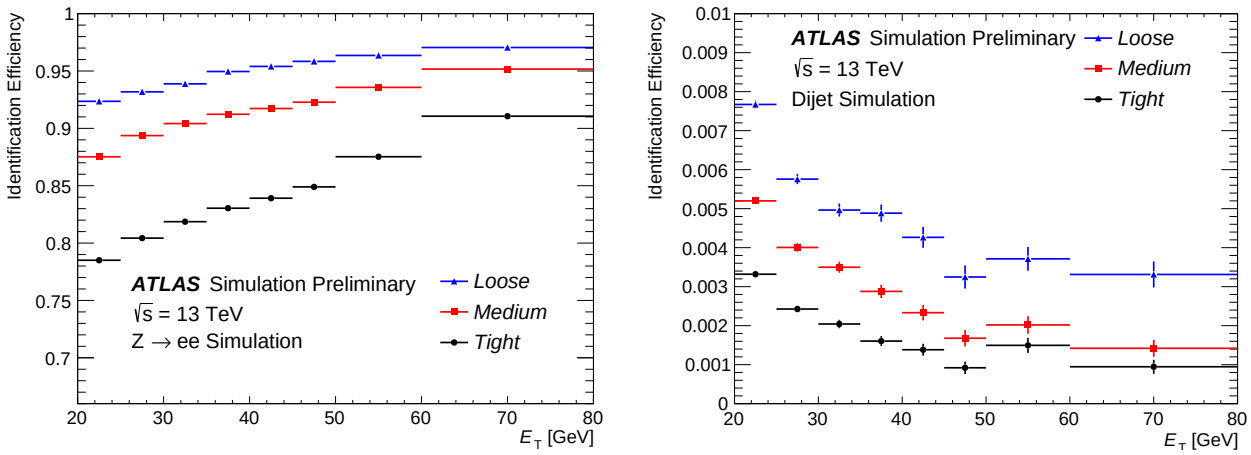
We apply further requirements to reject the backgrounds for the electron candidates such as hadronic jets or converted photons, . The identification algorithm is based on a MVA technique that simultaneously evaluates several properties of the electron candidates to make a decision. This method exploits quantities related to the electron cluster and track measurements including shower shapes in the calorimeter, measurements by the transition radiation tracker, quantities related to the matching between track and cluster, track properties, and variables measuring bremsstrahlung effects. These quantities are summarised in Table 4.1. We calculate an overall probability for the candidate from the probability distribution functions (PDFs) of these variables. This enables better background rejection for a given signal efficiency than a cut-based algorithm that uses selection criteria sequentially on each variable.

We use three levels of identification operating points, referred to as *Loose*, *Medium*, and *Tight*, increasing order of background rejection. The identification efficiency to the reconstructed electrons is estimated with the Tag and Probe method [119] using $Z^0 \to e^+e^-$ and its performance is shown in Figure 4.7. The signal efficiencies for electron candidates with $E_T = 25$ GeV are depending on the operating points, in the range from 78 to 90 % and increase with E_T; the background efficiencies are ranging from 0.3 to 0.8 % and decrease with E_T.

The electron identification performance can be influenced by the multiple collisions taking place in the same bunch crossing (in-time pileup) or in consecutive bunch crossings (out-of-time pileup). The number of reconstructed primary vertices indicates the level of pileup in the event. The distributions of the shower shower shape variables are affected by the number of pileup collisions per bunch crossing; $R_{\text{had}}, R_{\text{had1}}$ are the most influenced variables which are ratio between E_T in the hadronic calorimeter and in the electron cluster in the EM calorimeter, and are used to evaluate the hadronic leakage. Hence the cut on the discriminant value in MVA is optimised with the number of primary vertices, thereby the identification efficiency remains high without drastically increasing the amount of background accepted by the identification selection, as shown in Figure 4.8.

TABLE 4.1: Definitions of electron discriminating variables [119].

Type	Description	Name
Hadronic leakage	Ratio of E_{T} in the first layer of the hadronic calorimeter to E_{T} of the EM cluster (used over the range $\lvert\eta\rvert < 0.8$ or $\lvert\eta\rvert > 1.37$)	R_{had1}
	Ratio of E_{T} in the hadronic calorimeter to E_{T} of the EM cluster (used over the range $0.8 < \lvert\eta\rvert < 1.37$)	R_{had}
Back layer of EM calorimeter	Ratio of the energy in the back layer to the total energy in the EM accordion calorimeter. This variable is only used below 100 GeV because it is known to be inefficient at high energies.	f_3
Middle layer of EM calorimeter	Lateral shower width, $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	$w_{\eta 2}$
	Ratio of the energy in 3×3 cells over the energy in 3×7 cells centered at the electron cluster position	R_ϕ
	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position	R_η
Strip layer of EM calorimeter	Shower width, $\sqrt{(\Sigma E_i (i - i_{\mathrm{max}})^2)/(\Sigma E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	w_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
	Ratio of the energy in the strip layer to the total energy in the EM accordion calorimeter	f_1
Track conditions	Number of hits in the innermost pixel layer; discriminates against photon conversions	n_{Blayer}
	Number of hits in the pixel detector	n_{Pixel}
	Number of total hits in the pixel and SCT detectors	n_{Si}
	Transverse impact parameter with respect to the beam-line	d_0
	Significance of transverse impact parameter defined as the ratio of d_0 and its uncertainty	d_0/σ_{d_0}
	Momentum lost by the track between the perigee and the last measurement point divided by the original momentum	$\Delta p/p$
TRT	Likelihood probability based on transition radiation in the TRT	eProbabilityHT
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track	$\Delta\eta_1$
	$\Delta\phi$ between the cluster position in the middle layer and the track extrapolated from the perigee	$\Delta\phi_2$
	Defined as $\Delta\phi_2$, but the track momentum is rescaled to the cluster energy before extrapolating the track from the perigee to the middle layer of the calorimeter	$\Delta\phi_{\mathrm{res}}$
	Ratio of the cluster energy to the track momentum	E/p



(a) (b)

FIGURE 4.7: (a) The efficiency to identify electrons in $Z \rightarrow ee$ decays and (b) the efficiency to mis-identify hadrons as electrons estimated using simulated dijet samples. The efficiencies are obtained using $Z \rightarrow ee$ and dijet MC simulation samples, and are measured with respect to reconstructed electrons. The candidates are matched to true electrons for $Z \rightarrow ee$ events for (a). For background rejection studies the reconstructed electrons matched to the true electrons are not included [118].

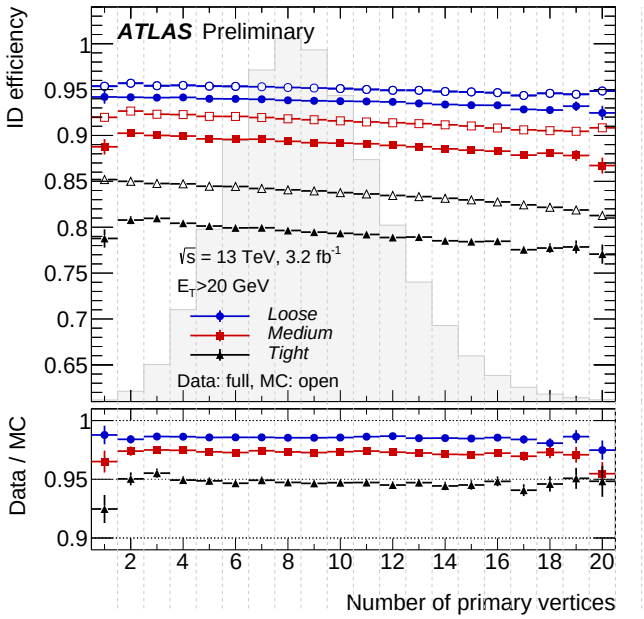
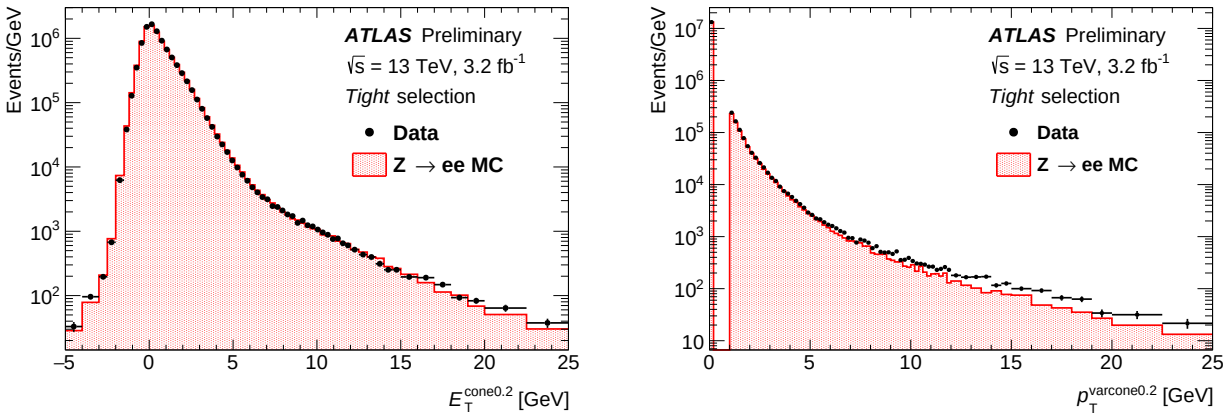


FIGURE 4.8: Electron identification efficiencies in the three operating points; Loose, Medium, and Tight, obtained using the Z Tag-and-Probe method, as a function of the number of reconstructed primary vertices. The total statistical and systematic uncertainties are shown. The distribution of the number of reconstructed primary vertices in the selected data events is overlaid in grey.



(a) (b)

FIGURE 4.9: Distributions of (a) $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ and (b) $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ for electrons in $Z \to ee$ events, in data (dots with error bars) and simulation (full histogram, normalised to data) [119].

4.2.5 Electron isolation

In addition to the identification criteria described above, we apply 'isolation' requirements to electrons, which allow us to distinguish prompt electrons (from heavy resonance decays to vector bosons, such as $W \to e\nu$, $Z \to ee$) from other non-isolated electron candidates such as electrons originating from converted photons produced in hadron decays, electrons from heavy flavour hadron decays, and light hadrons misidentified as electrons. We use two discriminating variables to quantify the energy of the particles produced around the electron candidate:

- calorimetric isolation variable; $E_{\mathrm{T}}^{\mathrm{cone0.2}}$, defined as the sum of transverse energies of clusters within a radius of $\Delta R = 0.2$ around the cluster for the candidate electron.

- track isolation variable; $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$, defined as the sum of transverse momenta of all tracks originating from the reconstructed primary vertex of hard collision, excluding the track associated to the electron. Here, the sum of p_{T} is calculated within a certain radius around the candidate electron track; we use $\Delta R = 0.2$ (10 GeV/E_{T}) for electrons of $E_{\mathrm{T}} < 50$ GeV ($E_{\mathrm{T}} > 50$ GeV).

The distributions of these two discriminating variables are shown in Figure 4.9 for electrons with E_{T} greater than 27 GeV satisfying the Tight identification requirements. The events are consistent between data and the $Z \to ee$ process. The negative tail of $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ originates from the correction for pileup and underlying event. A slight discrepancy is observed in the regions with large $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ and $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ values.

To select isolated electron candidates, we use two types of selection requirements on the quantities $E_{\mathrm{T}}^{\mathrm{cone0.2}}/E_{\mathrm{T}}$ and $p_{\mathrm{T}}^{\mathrm{varcone0.2}}/E_{\mathrm{T}}$:

- efficiency-targeted operating points with variable requirements to obtain a given isolation efficiency of $\epsilon_{\mathrm{iso}} =$ 90(99) % for $E_{\mathrm{T}} = 25(60)$ GeV.

- fixed-requirement operating points with the upper thresholds on the isolation variables.

The definitions of the operating points are listed in Table 4.2.

We use a fixed-requirement operating point ("*FixedCutTight*") to select signal electrons to reconstruct a resonance state, thereby we can use use low E_{T} electrons requiring high background rejection. We also use an efficiency-targeted operating point ("*LooseTrackOnly*") to reject the two-or-more lepton final states.

4.2.6 Electron trigger

In the online data process, we reconstruct electron trigger candidates at the Level-1 and the HLT stages (see section 2.2.4). The Level-1 electron triggers use both the EM and the hadronic calorimeter information within regions of $\Delta\eta \times \Delta\phi \approx 0.4 \times 0.4$ not only to reconstruct electron candidates above a certain energy threshold but also to apply

TABLE 4.2: Definitions of the electron isolation operating points. The upper table shows the efficiency-targeted operating points, and the numbers expressed in percents represent the target efficiencies used in the optimisation procedure of the operating points. For the Gradient and GradientLoose operating points, the unit of E_{T} is in GeV. The fixed-requirement operating points are shown in the lower table. The calorimeter and track isolation refer to the selection based on $E_{\mathrm{T}}^{\mathrm{cone0.2}}/E_{\mathrm{T}}$ and $p_{\mathrm{T}}^{\mathrm{varcone0.2}}/E_{\mathrm{T}}$, respectively [119].

	Efficiency		
Operating point	calorimeter isolation	track isolation	total efficiency
LooseTrackOnly	-	99%	99%
Loose	99%	99%	$\sim 98\%$
Tight	96%	99%	$\sim 95\%$
Gradient	$0.1143\% \times E_{\mathrm{T}} + 92.14\%$	$0.1143\% \times E_{\mathrm{T}} + 92.14\%$	90/99% at 25/60 GeV
GradientLoose	$0.057\% \times E_{\mathrm{T}} + 95.57\%$	$0.057\% \times E_{\mathrm{T}} + 95.57\%$	95/99% at 25/60 GeV

	Cut value	
Operating point	calorimeter isolation	track isolation
FixedCutLoose	0.20	0.15
FixedCutTightTrackOnly	-	0.06
FixedCutTight	0.06	0.06

isolation requirements. At the HLT stage, electron trigger candidates are reconstructed by an offline-like algorithm using information from calorimeter and tracking. The identification (ID) algorithm in the HLT stage is performed via the same discriminating variables as the offline identification described in subsection 4.2.4. The identification criteria are slightly looser than the offline requirements to avoid efficiency losses. Two types of approaches exist for the identification criteria; (1) the cut-based approach and (2) the likelihood-based approach [119].

During LHC Run-1, the electron trigger with the cut-based approach was used. However, inefficiency can arise by applying the likelihood-based identification offline with a cut-based identification in the trigger, thus we also use the electron likelihood identification in LHC Run-2.

4.2.7 Efficiency measurement

The efficiency to find and select an electron in the ATLAS detector is not measured as a single quantity but is divided into different components, namely reconstruction, identification, isolation, and trigger efficiencies. The total efficiency $\epsilon_{\mathrm{total}}$ for a single electron can be written as:

$$\epsilon_{\mathrm{total}} = \epsilon_{\mathrm{reconstruction}} \times \epsilon_{\mathrm{identification}} \times \epsilon_{\mathrm{isolation}} \times \epsilon_{\mathrm{trigger}} \quad , \tag{4.5}$$

with the various efficiency components are determined in the tag-and-probe method. Combined efficiencies from the reconstruction and identification are shown in Figure 4.10 and the isolation and trigger efficiencies are shown in Figure 4.11. Since the electron efficiencies depend on E_{T} and η, the measurements are performed in two dimensional bins in (E_{T}, η).

4.3 Muon reconstruction

Muon candidates are first reconstructed independently in the inner detector (ID) and the muon spectrometer (MS). The information from each of the sub-detectors is then combined to form muon tracks to be used in physics analyses. After several corrections on the muon momentum scale and resolution between the data and MC, we select muons based on track parameters. We will overview each of the steps in this section.

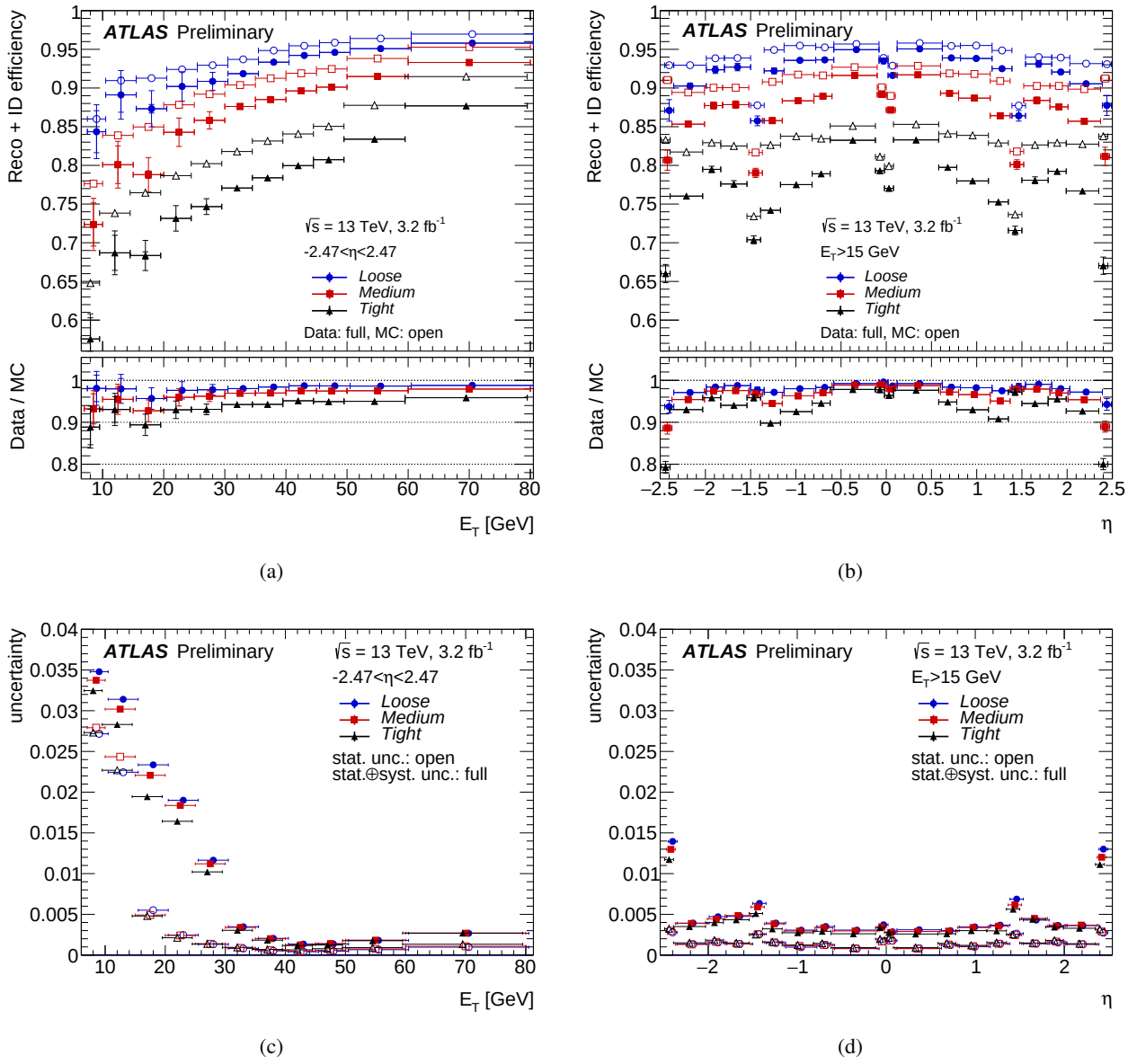
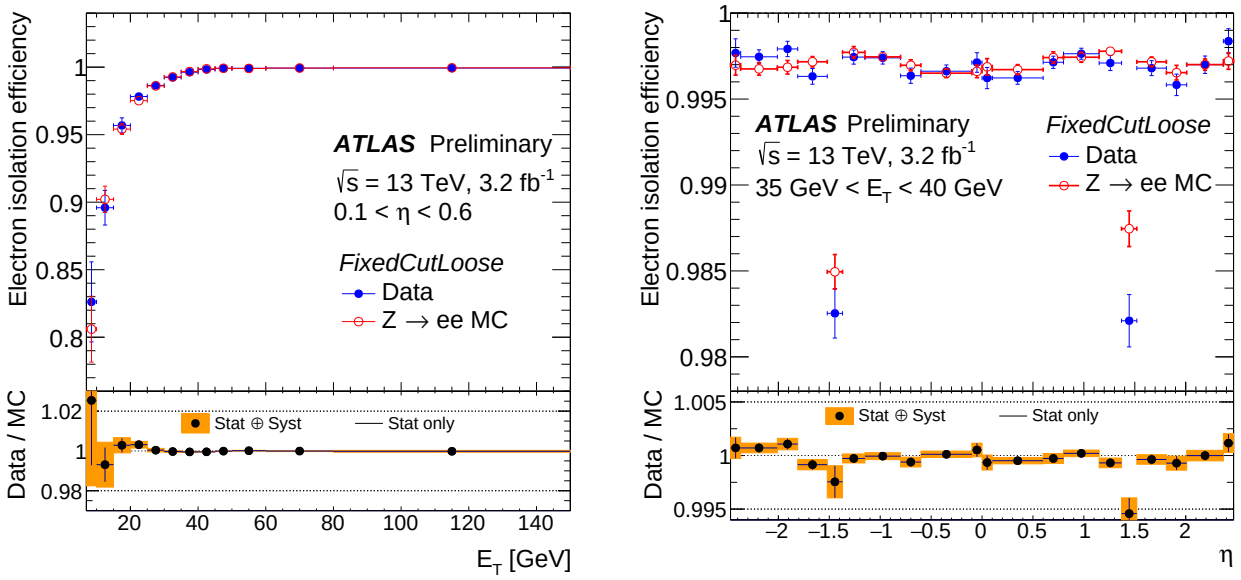
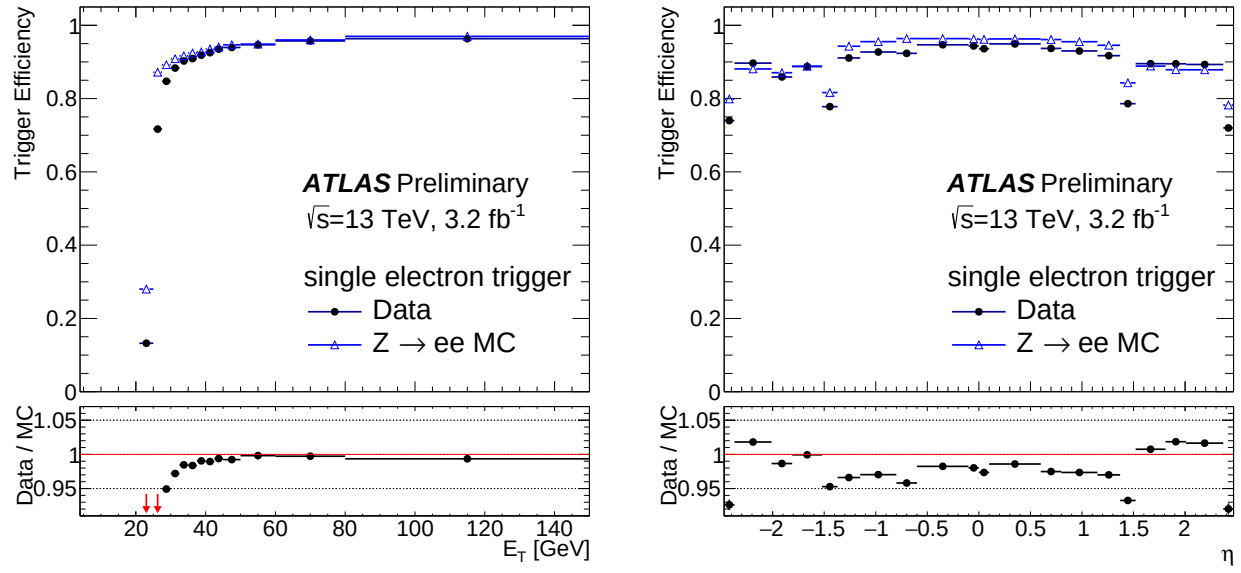


FIGURE 4.10: Electron reconstruction, identification combined efficiency (a) as a function of E_{T}, and (b) as a function of η; its absolute uncertainties in $Z \to ee$ events (c) as a function of E_{T}, and (d) as a function of η [119].



(a) (b)



(c) (d)

FIGURE 4.11: Efficiency of the FixedCutLoose isolation requirement ($E_{\mathrm{T}}^{\mathrm{cone0.2}}/E_{\mathrm{T}} < 0.2$ and $p_{\mathrm{T}}^{\mathrm{varcone0.2}}/E_{\mathrm{T}} < 0.15$) for electrons in $Z \to ee$ (a) as a function of E_{T} and (b) η; Trigger efficiency for the single electron trigger measured in $Z \to ee$ events as a function of (c) E_{T} and (d) η [119].

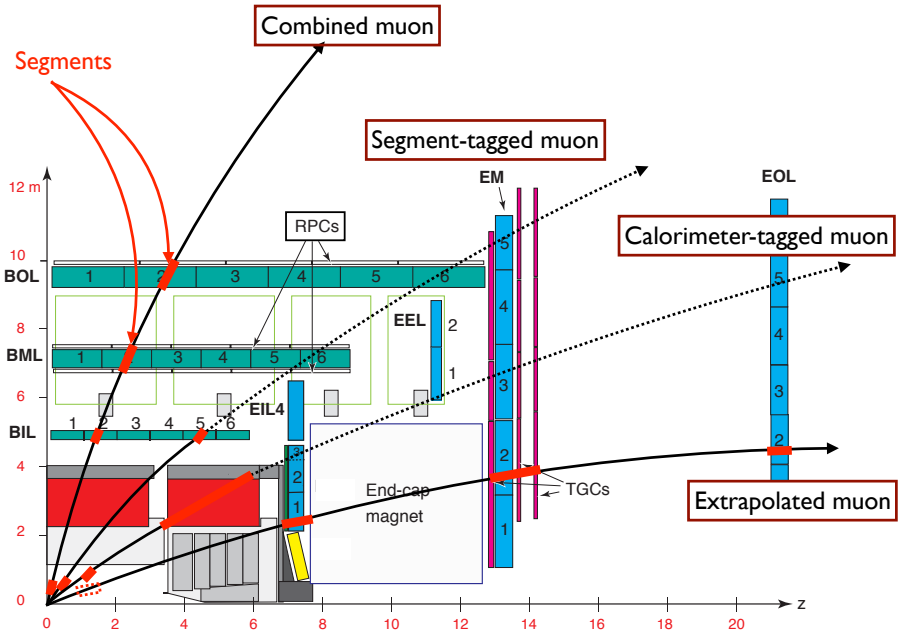


FIGURE 4.12: Schematic view of the muon reconstruction. Combined, segment-tagged, calorimeter-tagged, and extrapolated muons are defined in the main body of this sections.

4.3.1 Muon reconstruction in Muon Spectrometer

Muon reconstruction begins in the MS by searching for hit patterns on each of the muon detectors (MDT, CSC, RPC, and TGC) to form 'segments'. We use a Hough transform [120] in each of the MDT chambers and nearby trigger chambers to search hits aligned on a trajectory on the bending plane. The MDT segments are reconstructed by performing a straight-line fit to the hits found in each layer; the RPC (for the barrel region) or TGC (for the end-cap region) hits measure the coordinate orthogonal to the bending plane. In CSC a different algorithm is used; segments in CSC are built using a separate combinatorial search in the η and ϕ planes. Then we build muon track candidates with fitting hits from segments in the muon detectors. The fitting algorithm starts by using the segments in the middle layers and then the fitting is extended to the segments in the inner and outer layers. The reconstructed tracks are extrapolated to the IP, taking into account the energy loss at the calorimeter.

4.3.2 Combined reconstruction

Tracks reconstructed in the ID can help to improve the precision of the muon tracks and suppress mis-identification of the in-flight decays of pions and kaons. The combination between ID and MS information is performed by the following four algorithms, which are also illustrated in Figure 4.12;

- Combined muons: muon trajectories independently reconstructed in the ID and MS are combined by a global fit. An outside-in pattern recognition is employed, in which the muons are first reconstructed in the MS and then extrapolated to an ID track. The track combination is performed only for pairs that show a matching in the $\eta - \phi$ plane at the closest approach to the beam line.
- Segment-tagged muons: tracks in the ID are associated with at least one local segments in the MDT or CSC chambers.
- Calorimeter-tagged muons: a track in the ID is identified as muon if it matches an energy deposit in the calorimeter corresponding to a minimum-ionising particle.
- Extrapolated muons: muon trajectories are reconstructed based on the MS tracks applying a loose requirement that the trajectory originate from the IP.

A muon track can be recognised in multiple types of the reconstruction algorithms. When two muon candidates share the same ID track, combined muons are in higher priority, then segment-tagged muons, and finally calorimeter-tagged muons. The overlap with extrapolated muons is resolved by selecting a track with better fit quality and larger number of hits.

4.3.3 Muon identification

To suppress background effects, muon identification is performed by applying a set of quality selections. We have selections; *Loose*, *Medium*, *Tight*, and *High-*p_{T}, increasing order of background rejection. Loose, Medium and Tight selections are inclusive selections with tighter requirements.

- Loose muons: The Loose muon criteria is provided to maximise the reconstruction efficiency which slightly larger uncertainty than the Medium muons.
- Medium muons: This is the default criteria for muons in ATLAS and minimise the systematic uncertainties associated with muon reconstruction and calibration.
- Tight muons: The Tight muons are adopted in some analyses to require the higher purity of muons than the Medium muons at the expense of some efficiency.
- High-p_{T} muons: The High-p_{T} muons minimise the momentum resolution for tracks with $p_{\mathrm{T}} > 100$ GeV.

We use the Medium requirement to select signal muons to reconstruct a resonance state, and the Loose identification to reject two or more lepton final states.

The muon reconstruction efficiencies are evaluated by a Tag-and-probe method [121] based on the selections of almost pure muon tracks in $J/\Psi \to \mu\mu$ in low p_{T} region both in data and simulation samples. In order to evaluate the efficiencies in the high p_{T} region, we exploit the same method using $Z \to \mu\mu$ events. Figure 4.13 shows the reconstruction efficiency and its uncertainties binned in η and p_{T}. The disagreement of measured efficiency between data (ϵ^{data}) and simulation samples (ϵ^{MC}) is expressed as the "scale factor" (SF):

$$\mathrm{SF} = \frac{\epsilon^{\mathrm{data}}}{\epsilon^{\mathrm{MC}}} \quad , \tag{4.6}$$

This quantity describes the difference between simulation and real data, and is used to correct the simulation samples.

4.3.4 Muon isolation

We exploit muons originating from the decay of W and Z bosons in the analysis. These muons are well separated from other particles, unlike muons in jets, thus we can select the W/Z-originated muons by applying further isolation selections. Similar to the electron isolation, seven isolation criteria described in Table 4.3 are defined based on the muon isolation variables:

- track-based isolation variable; $p_{\mathrm{T}}^{\mathrm{varcone3.0}}$, which is defined as the scalar sum of the transverse momenta of the tracks within a certain radius of $\Delta R = 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu}$ for muons with $p_{\mathrm{T}} > 33$GeV (At $p_{\mathrm{T}} = 33$GeV, $\Delta R = 0.3$) around the muon p_{T}^{μ} except for muon track itself. For muons with lower p_{T} ($p_{\mathrm{T}} < 33$GeV), a flat radius of ΔR value of 0.3 is adopted.
- calorimeter-based isolation variable; $E_{\mathrm{T}}^{\mathrm{topocone2.0}}$, which is defined as the sum of the transverse energy of "topo-clusters", which is described in Section 4.4.1, within a radius of $\Delta R = 0.2$ around the muon, subtracting the energy deposit by the muon itself and mitigating the pile-up effects.

We use the FixedCutTightTrackOnly requirement to define signal muons to reconstruct a resonance state, and the LooseTrackOnly requirement to reject two or more lepton final states. Isolation efficiencies for the LooseTrackOnly and the FixedCutLoose working points are shown in Figure 4.14 binned in p_{T}.

4.3.5 Momentum scale calibration

To correct the small mis-modelling of the muon momentum scale due to the incomplete description of the magnetic field, we apply the muon momentum scale calibration algorithm defined in this section. The corrected muon transverse momentum, $p_{\mathrm{T}}^{\mathrm{Cor,Det}}$ [121], is described in:

$$p_{\mathrm{T}}^{\mathrm{Cor,Det}} = \frac{p_{\mathrm{T}}^{\mathrm{MC,Det}} + \sum_{n=0}^{1} s_n^{\mathrm{Det}}(\eta,\phi)(p_{\mathrm{T}}^{\mathrm{MC,Det}})^n}{1 + \sum_{m=0}^{2} \Delta r_m^{\mathrm{Det}}(\eta,\phi)(p_{\mathrm{T}}^{\mathrm{MC,Det}})^{m-1} g_m} \quad , \tag{4.7}$$

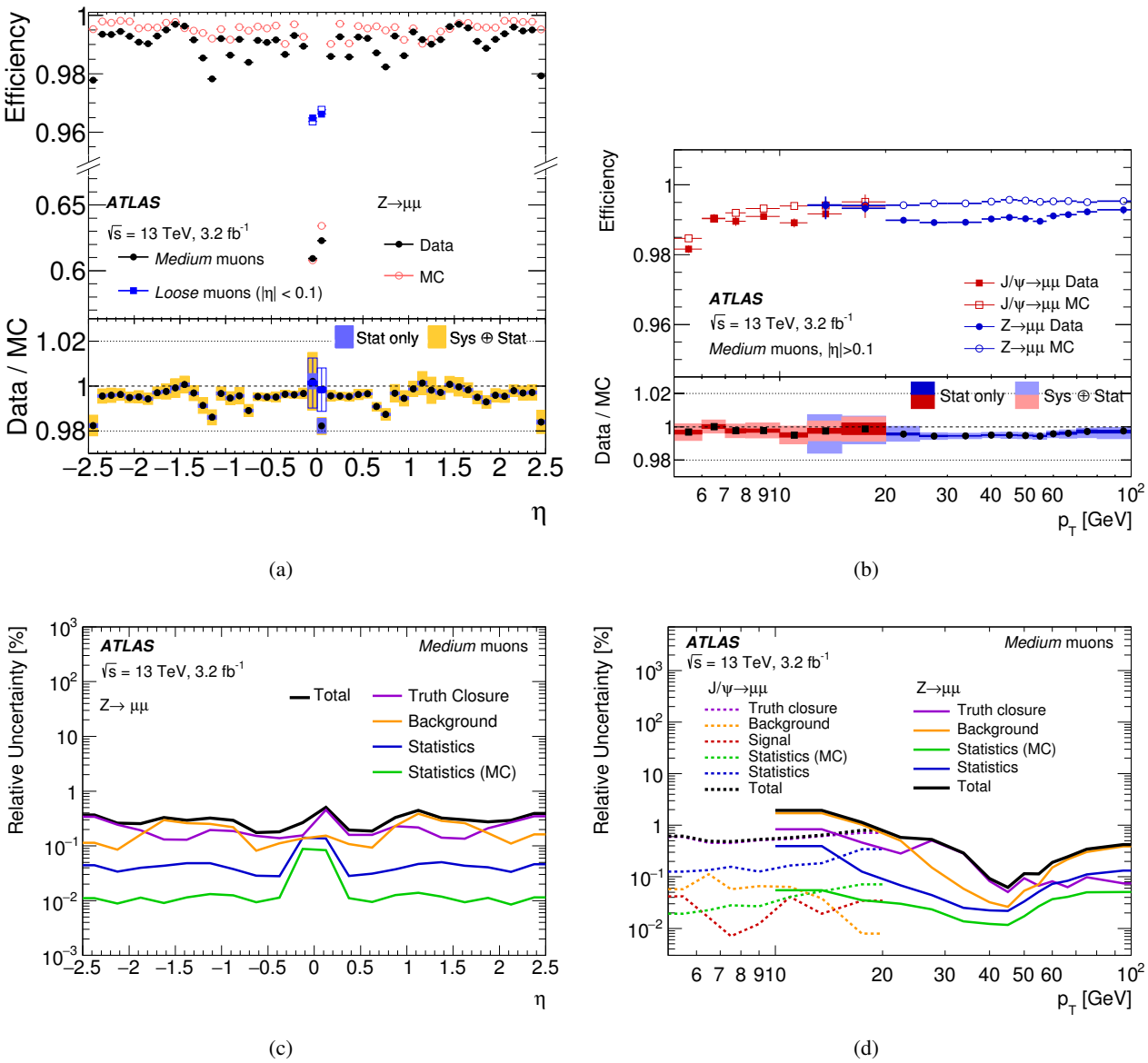
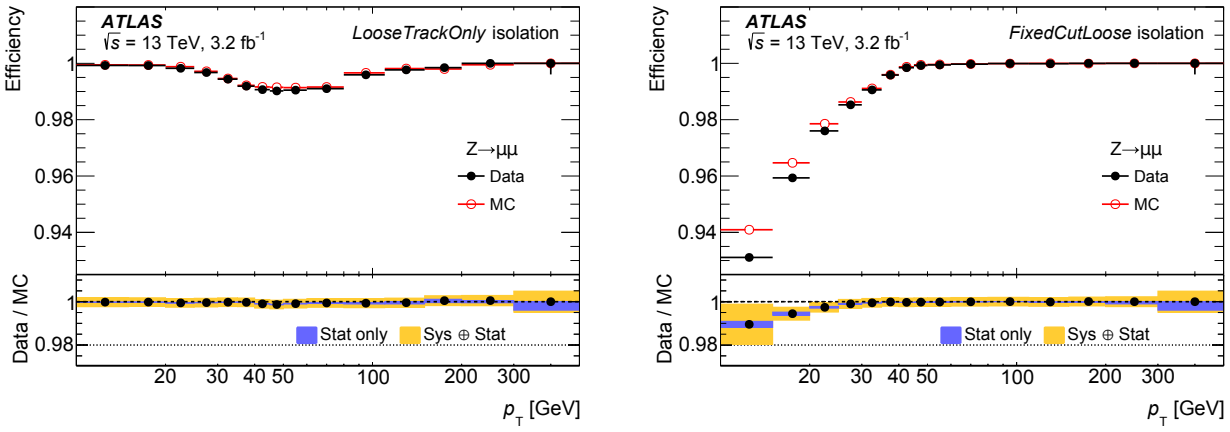


FIGURE 4.13: Muon reconstruction efficiency (a) binned in η and (b) binned in p_{T}; its uncertainties (c) binned in η, and (d) binned in p_{T} [121]. The efficiencies and uncertainties are evaluated by a Tag-and-probe method using $J/\Psi \to \mu\mu$ ($Z \to \mu\mu$) events in low (high) p_{T} regions.

TABLE 4.3: Definition of the seven isolation working points for muons [121]. The discriminating variables are listed in the second column and the criteria used in the definition are described in the third column.

Isolation WP	Discriminating variable(s)	Definition
LooseTrackOnly	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$	99% efficiency constant in η and p_{T}
Loose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	99% efficiency constant in η and p_{T}
Tight	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	96% efficiency constant in η and p_{T}
Gradient	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	$\geq$ 90(99)% efficiency at 25 (60) GeV
GradientLoose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	$\geq$ 95(99)% efficiency at 25 (60) GeV
FixedCutTightTrackOnly	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.06$
FixedCutLoose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.15$, $E_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu} < 0.30$



(a) (b)

FIGURE 4.14: Isolation efficiency (a) with the LooseTrackOnly working point and (b) with the FixedCutLoose working point. [121] The efficiency is shown as a function of the muon p_{T} and is measured in $Z \to \mu\mu$ events. The error bars on the top panel are statistical only, while the combination of statistical and systematic uncertainties is shown on the bottom panel as well.

where $p_{\mathrm{T}}^{\mathrm{MC,Det}}$ is the uncalibrated transverse momentum in simulation (Det=MS or ID), g_m are normally distributed random variables with mean 0 and width 1, and the terms $\Delta r_m^{\mathrm{Det}}(\eta,\phi)$ and $s_n^{\mathrm{Det}}(\eta,\phi)$ describe the momentum resolution and the scale corrections applied in a specific η and ϕ region, respectively. Δr_m corrects each term of:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{r_0}{p_{\mathrm{T}}} \oplus r_1 \oplus r_2 \cdot p_{\mathrm{T}} \quad , \tag{4.8}$$

where $\oplus$ denotes sum in quadrature. The first term accounts for fluctuations of the energy loss in the traversed material, the second term accounts for the multiple scattering, inhomogeneities of the local magnetic field, and local radial displacements, and the third term accounts for intrinsic resolution effects caused by the spatial resolution of the hit measurements and residual misalignment. The energy loss term is negligible both in the ID and MS, therefore Δr_0^{ID} and Δr_0^{MS} are set to zero.

The invariant mass distributions of the $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ candidates are shown in Figures 4.15, in which both uncalibrated and calibrated simulation samples are shown. In the uncalibrated simulation, it is noticeable that the signal distributions are narrower and slightly shifted with respect to data. After the calibration, the peak shape of the two resonances in simulation agree with the data within the systematic uncertainties, demonstrating the overall effectiveness of the p_{T} calibration.

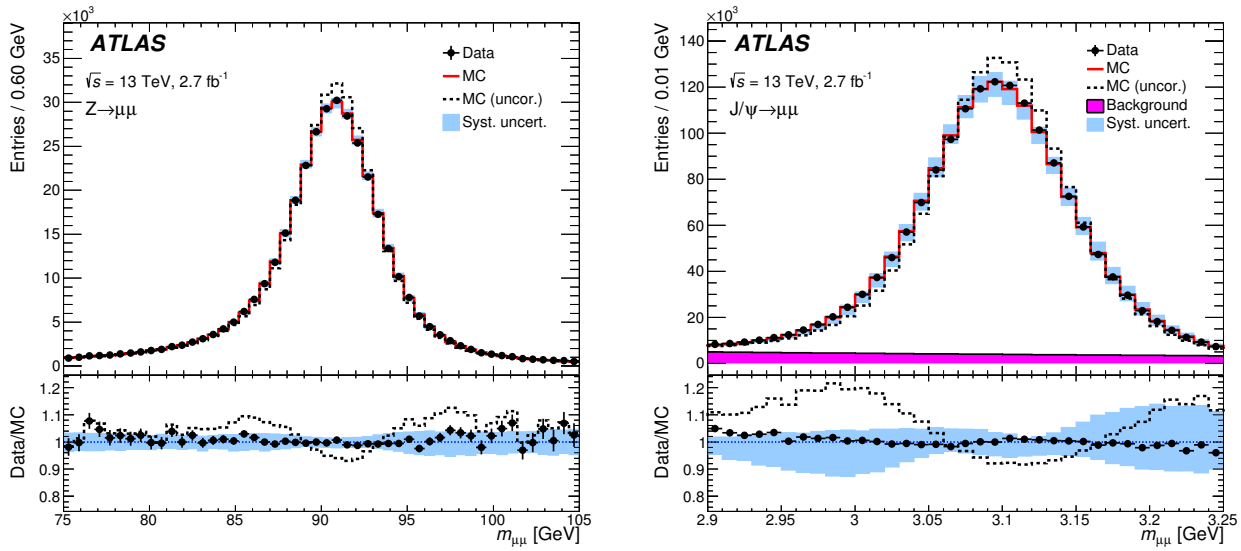
4.4 Jet reconstruction

We reconstruct collimated showers of hadrons, namely jets, based on "topo-clusters" reconstructed in the calorimeter. We describe a overview of a jet reconstruction algorithm at first, then mitigation-method for pile-up effects and tagging algorithms for b-quark induced jets. Since the performance of the jet energy calibration plays an important role for the resonance search in this thesis, the jet-calibration flow is described in Chapter 5.

4.4.1 Topo-cluster reconstruction

We form a topological cluster ("topo-cluster") as a combination of energies in the calorimeter cells to find significant patterns generated by particle showers. The topo-cluster definition relies on the significance measured in the calorimeter cells, $\varsigma_{\mathrm{cell}}^{\mathrm{EM}}$ [122], defines as;

$$\varsigma_{\mathrm{cell}}^{\mathrm{EM}} = \left| \frac{E_{\mathrm{cell}}^{\mathrm{EM}}}{\sigma_{\mathrm{noise,cell}}^{\mathrm{EM}}} \right| \quad , \tag{4.9}$$



(a) (b)

FIGURE 4.15: Dimuon invariant mass distribution of (a) $Z \to \mu\mu$ and (b) $J/\Psi \to \mu\mu$ candidates reconstructed with combined muons [121]. The upper panels show the invariant mass distributions for data and for the signal simulation plus the background estimate.The points show the data. The continuous line corresponds to the simulation with the momentum corrections applied and the dashed lines show the simulation when no correction is applied.

where $E^{\mathrm{EM}}_{\mathrm{cell}}$ is the cell energy and $\sigma^{\mathrm{EM}}_{\mathrm{noise,cell}}$ is the expected cell noise. The topo-cluster reconstruction algorithm, which is illustrated in Figure 4.16, begins by forming *proto-clusters* with the initial cells requiring them to have $\varsigma^{\mathrm{EM}}_{\mathrm{cell}} > 4$. Then, we collect neighbouring cells with $\varsigma^{\mathrm{EM}}_{\mathrm{cell}} > 2$ for the proto-clusters. If two proto-clusters contain the same cell with absolute energy 2σ above the noise threshold, these proto-clusters are merged. After all neighbouring cells with absolute energies above 2σ have been collected, all the neighbouring cells ($\varsigma^{\mathrm{EM}}_{\mathrm{cell}} \geq 0$) are added to the cluster. This set of thresholds is commonly known as "4-2-0" topo-cluster reconstruction.

After the proto-cluster formation, any proto-clusters with two or more local maxima, which is signified by $E^{\mathrm{EM}}_{\mathrm{cell}} > 500$ MeV, can be split into separate clusters. In addition, following requirements are applied to the proto-clusters; the maximum cell has at least four neighbouring cells, and none of the neighbouring cells have a larger signal. In principle, cells can be shared between two proto-clusters, with the cell assigned a fractional weight for each cluster based on the energies of the clusters, and the distances of the cell to the centre of gravity of the two clusters (d_1, d_2, here $d_1 < d_2$). The fractional weight is given by [122];

$$\mathcal{W}_{\mathrm{cell},1} = \frac{E^{\mathrm{EM}}_{\mathrm{clus},1}}{E^{\mathrm{EM}}_{\mathrm{clus},1} + rE^{\mathrm{EM}}_{\mathrm{clus},2}}, \quad \mathcal{W}_{\mathrm{cell},2} = 1 - \mathcal{W}_{\mathrm{cell},1} \quad , \tag{4.10}$$

where $r = \exp(d_1 - d_2)$.

4.4.2 anti-k_t algorithm

After the topo-cluster reconstruction, we reconstruct hadronic jets from the reconstructed clusters in the 'anti-k_t' algorithm [34]. There are several jet clustering algorithms, such as k_t [123] and Cambridge/Aachen algorithm [124], and the anti-k_t algorithm, in which we commonly introduce a distance variable: d_{ij}; distance between two clusters i and j, and d_{iB} is distance between cluster i and the beam (B):

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2} \quad , \tag{4.11}$$

$$d_{iB} = k_{ti}^{2p} \quad , \tag{4.12}$$

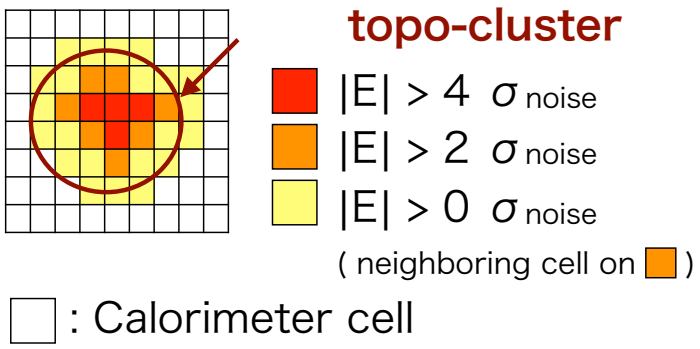
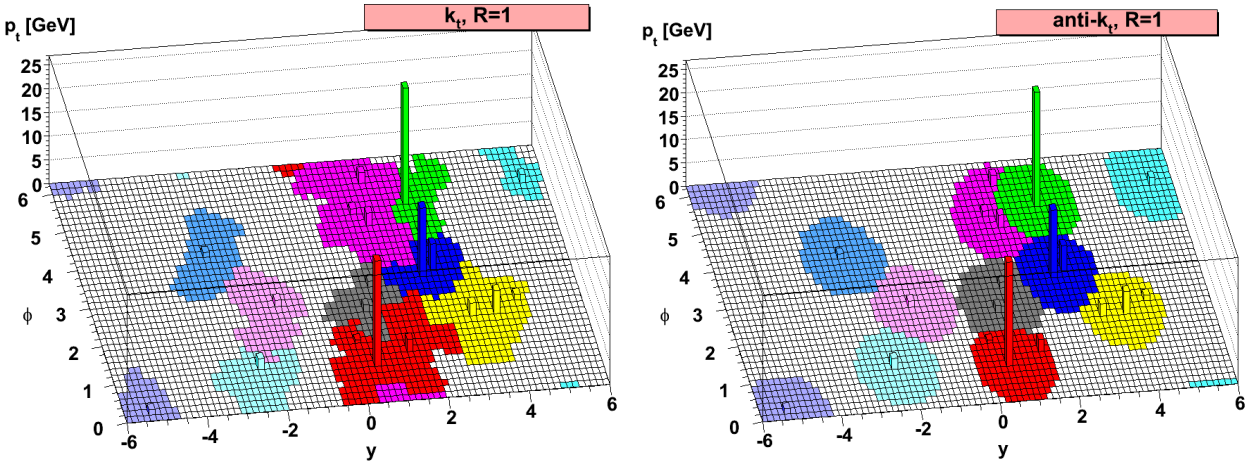


FIGURE 4.16: Schematic view of the formation of topo-cluster.



(a) (b)

FIGURE 4.17: An example of a parton-level event, together with many random soft components, clustered with (a) k_t algorithm and (b) anti-k_t algorithm [34]. For both cases, the distance parameter R is 1. Hadron jets reconstructed in the anti-k_t algorithm are more clear than those in the k_t algorithm.

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i, and ϕ_i are the transverse momentum, rapidity, and azimuth angle of cluster i, respectively. R denotes the radius parameter, which represents the cone size of jets, approximately. In Equation 4.11 and 4.12, we have added a parameter p to govern the relative power of the energy versus geometrical (Δ_{ij}) scales.

We can cover k_t (Cambridge/Aachen) algorithm by choosing $p = 1$ ($p = 0$), however we use the anti-k_t algorithm with $p = -1$. In the anti-k_t algorithm, the distance variables d_{ij} and d_{iB} are computed with all the topo-clusters. If the smallest parameter among them is a d_{ij}, we combine two topo-clusters i and j. If the smallest parameter among all distances is a d_{iB}, we remove the i-th topo-cluster from the list and adopt it as a jet. The distance parameters are re-calculated and the procedure is repeated until all topo-clusters are clustered into jets.

Figure 4.17 shows examples of the jet-clustering in the k_t and the anti-k_t algorithms. We can reconstruct conical jets by the anti-k_t algorithm; the reason of this feature is explained in Figure 4.18. In the anti-k_t algorithm, $d_{1i} = \min(1/k_{t1}^2, 1/k_{ti}^2)\Delta_{1i}^2/R^2$ between a hard cluster 1 and a soft cluster i is determined by the transverse momentum of the hard cluster and the Δ_{1i} separation. d_{ij} between soft clusters i and j with a similar separation $\Delta_{ij} \sim \Delta_{1i}$ is much larger than d_{1i}. Therefore soft clusters are likely to be combined with the hard cluster before it is combined to other soft clusters. If a hard cluster has no hard neighbouring clusters within a distance $2R$, this cluster accumulate all the soft clusters within a circle of radius R. In the ATLAS experiment, the default choice of the jet reconstruction is the anti-k_t algorithm with $R = 0.4$ are default jets (small-R jets). In MC simulation samples, truth jets are also reconstructed in the anti-k_T algorithm with quasi stable particles except for muons and neutrinos.

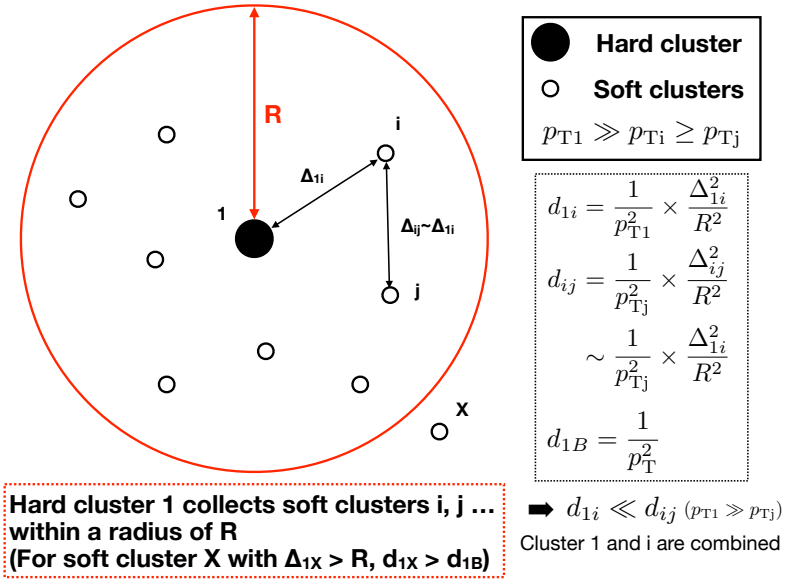


FIGURE 4.18: Schematic view of jet reconstruction by the anti-k_t algorithm.

4.4.3 Jet Vertex Tagging

To reject jets induced by pile-up, we use the jet vertex tagger (JVT) algorithm [125], which is constructed using two discriminants R_{pT} and corrJVF to separate hard-scatter events from pile-up, defined as:

$$R_{\mathrm{pT}} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk_k}}(\mathrm{PV}_0)}{p_{\mathrm{T}}^{\mathrm{jet}}} , \tag{4.13}$$

$$\mathrm{corrJVF} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk_k}}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T}}^{\mathrm{trk_l}}(\mathrm{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_{\mathrm{T}}^{\mathrm{trk_l}}(\mathrm{PV_n})}{k \cdot n_{\mathrm{trk}}^{\mathrm{PU}}}} , \tag{4.14}$$

where $\sum_k p_{\mathrm{T}}^{\mathrm{trk_k}}(\mathrm{PV}_0)$ is scalar sum of the track p_{T} that are associated with the jet and originate from the hard-scatter vertex. The term $\sum_{n\geq 1}\sum_l p_{\mathrm{T}}^{\mathrm{trk_l}}(\mathrm{PV_n}) \equiv p_{\mathrm{T}}^{\mathrm{UE}}$ denotes scalar p_{T} sum of associated tracks that originate from any of the pileup interactions. The term $\langle p_{\mathrm{T}}^{\mathrm{UE}} \rangle$ linearly increase with the total number of pileup tracks per event $n_{\mathrm{trk}}^{\mathrm{PU}}$; to correct this effect, we divide $p_{\mathrm{T}}^{\mathrm{UE}}$ by $(k \cdot n_{\mathrm{trk}}^{\mathrm{PU}})$ with $k = 0.01$.

Figure 4.19 shows the correlation of corrJVF and R_{pT} for hard-scatter and pileup jets. From R_{pT} and corrJVF, the 2-dimensional likelihood, JVT, is defined as shown in Figure 4.19(c). The distribution of JVT is shown in Figure 4.19(d). In the resonance search presented in this thesis, jets with JVT< 0.59 are rejected in the region of $|\eta| < 2.4$ and $p_{\mathrm{T}} < 60$ GeV. Using the JVT selection, efficiency for the jets in the signal events is 92 % with pile-up fake rates of $\sim$ 2 %.

4.4.4 b-jet tagging

To distinguish WW/WZ events from $t\bar{t}$ events, we need to tag jets induced by b-quarks. Due to the long lifetime of hadrons containing a b-quark ($\sim$ 1.5 ps), tracks in decay products from b-hadrons have large impact parameters, and we can reconstruct a displaced secondary vertex with the jet. There are several algorithms which achieve b-jet tagging, relying on the impact parameter, or the secondary vertex finding algorithms [126]. In this thesis, however, we employ a BDT-based multivariate algorithm, namely MV2 algorithm [127], to achieve a better performance than any of the individual algorithms.

The input variables for the MV2 algorithm are listed in the Table 4.4. Using these input variables, three MV2 algorithms are developed; MV2c00, MV2c10, and MV2c20. The performance of each of the algorithms are shown in Figure 4.20. We adopt the MV2c10 algorithm in our analysis, because it provides the best rejection for the light-flavour jets of 4% at a working point with 77% of b-jet efficiency.

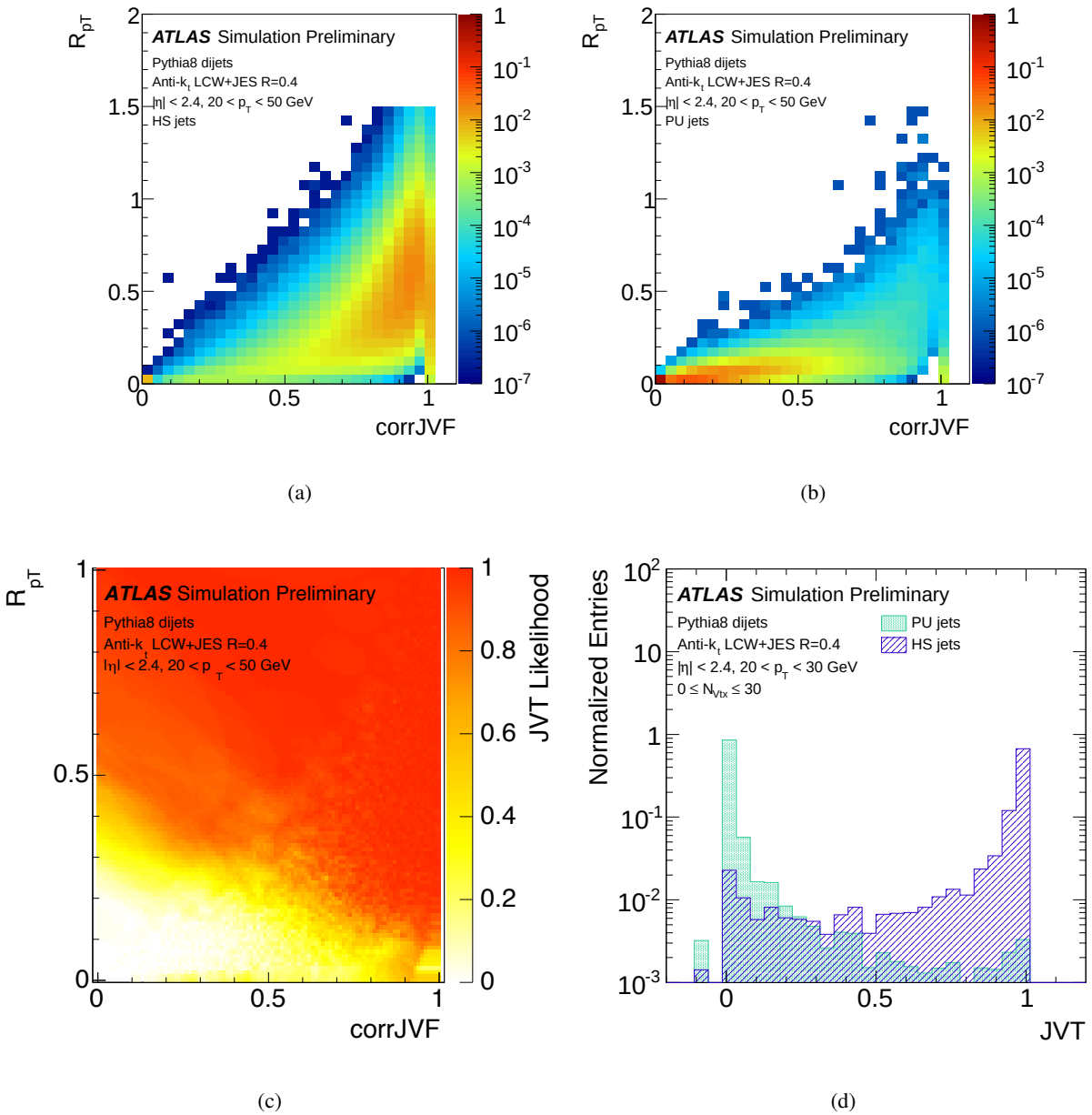
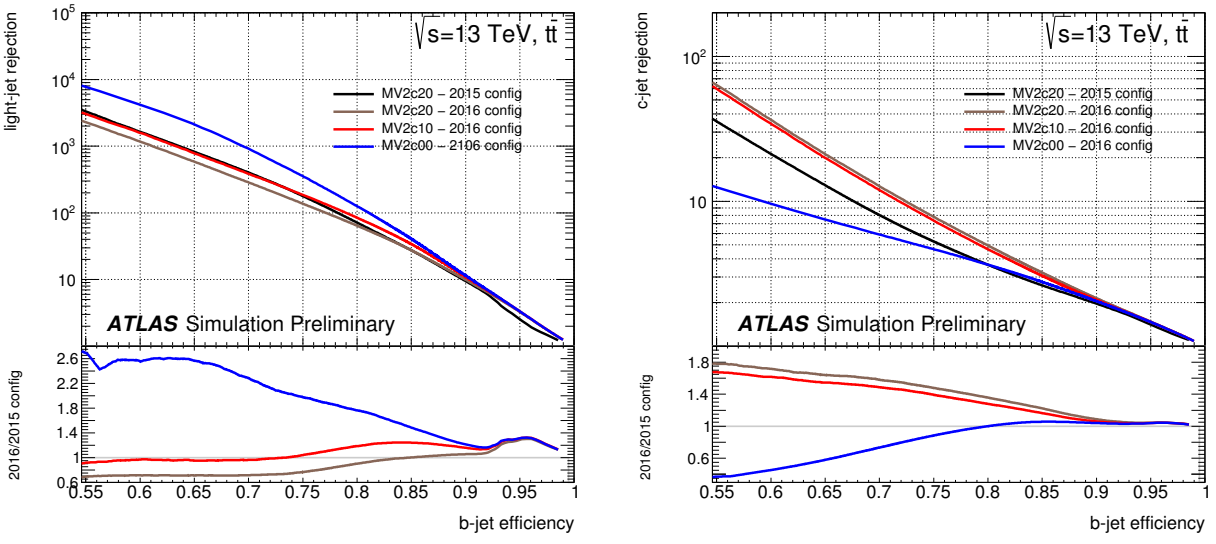


FIGURE 4.19: 2-dimensional correlation of corrJVF and R_{pT} for (a) hard-scatter and (b) pileup jets. (c) The 2-dimensional JVT likelihood as a function of corrJVF and R_{pT}. (d): Distribution of JVT for pile-up and hard0scatter jets with $20 < p_{\mathrm{T}} < 30$ GeV [125].

TABLE 4.4: The 24 input variables used by the MV2 algorithm [127].

Input	Variable	Description
Kinematics	$p_T(jet)$	Jet transverse momentum
	$\eta(jet)$	Jet pseudo-rapidity
IP2D, IP3D	$\log(P_b/P_{\text{light}})$	Likelihood ratio between the b- and light jet hypotheses
	$\log(P_b/P_c)$	Likelihood ratio between the b- and c-jet hypotheses
	$\log(P_c/P_{\text{light}})$	Likelihood ratio between the c- and light jet hypotheses
SV	$m(\text{SV})$	Invariant mass of tracks at the secondary vertex assuming pion masses
	$f_E(\text{SV})$	Fraction of the charged jet energy in the secondary vertex
	$N_{\text{TrkAtVtx}}(\text{SV})$	Number of tracks used in the secondary vertex
	$N_{\text{2TrkVtx}}(\text{SV})$	Number of two track vertex candidates
	$L_{xy}(\text{SV})$	Transverse distance between the primary and secondary vertices
	$L_{xyz}(\text{SV})$	Distance between the primary and secondary vertices
	$S_{xyz}(\text{SV})$	Distance between the primary and secondary vertices divided by its uncertainty
	$\Delta R(\text{jet}, \text{SV})$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex
Jet Fitter	$N_{\text{2TrkVtx}}(\text{JF})$	Number of 2-track vertex candidates (prior to decay chain fit)
	$m(\text{JF})$	Invariant mass of tracks from displaced vertices assuming pion masses
	$S_{\text{xyz}}(\text{JF})$	Significance of the average distance between the primary and displaced vertices
	$f_E(\text{JF})$	Fraction of the charged jet energy in the secondary vertices
	$N_{\text{1-trk vertices}}(\text{JF})$	Number of displaced vertices with one track
	$N_{\geq\text{2-trk vertices}}(\text{JF})$	Number of displaced vertices with more than one track
	$N_{\text{TrkAtVtx}}(\text{JF})$	Number of tracks from displaced vertices with at least two tracks
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})$	ΔR between the jet axis and the vectorial sum of the momenta of all tracks attached to displaced vertices



(a) (b)

FIGURE 4.20: Performance of the MV2 b-tagging algorithm evaluated on $t\bar{t}$ events [126]. (a) Light flavour jet and (b) c-jet rejection versus b-jet efficiency with the configurations in 2015 and 2016 data.

4.5 Large-R jet reconstruction

It is difficult to reconstruct separate jets by using the default jet reconstruction algorithm (anti-k_t $R = 0.4$) for highly collimated quarks. In our analysis, we reconstruct the vector bosons from high-mass resonance states are highly-boosted; the opening angles between the quarks from hadronic decay of the W/Z bosons are very small. In order to reconstruct those highly-boosted W/Z bosons, we exploit the 'large-R jets' techniques, i.e. we treat the two quarks as a single jet. The large-R jets are reconstructed in the anti-k_t algorithm with the large radius parameter $R = 1.0$. At first, we describe the large-R jet reconstruction method and a pile-up mitigation technique (called *grooming*). Then, we see boson tagging techniques and their performance.

4.5.1 Large-R jet reconstruction and grooming

To reconstruct large-R jets, topo-clusters individually calibrated for hadronic signals (local hadronic cell weighting, LCW [122]) are used to correct energies for effects such as the non-compensation of the calorimeter response, excitation of atoms, escaping neutrons, and inactive materials. We reconstruct large-R jets with the LCW topo-clusters using the anti-k_t algorithm with a radius parameter of $R = 1.0$.

Reconstructing jets with larger area induces more effects of pile-up and underlying events. To mitigate these effects, grooming techniques [128] are adopted for the large-R jets. Whereas there are several grooming techniques, such as 'filtering' [129] and 'pruning' [130], we use the 'trimming' [131] in which the original constituents of the jets are 'reclustered' using the k_t algorithm with a small distance parameter R_{sub} in order to reconstruct a collection of 'subjets'. These subjets are checked whether they carry p_{T} less than a specific fraction f_{cut} of the p_{T} of the original jet. The trimming parameters used for the analysis are $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 5\%$. A schematic view of the jet trimming procedure is shown in Figure 4.21(a). Figures 4.21(b) and 4.21(c) show the dependence of the mean jet mass, $\langle m^{\text{jet}} \rangle$, for trimmed and un-trimmed jets, as a function of number of the reconstructed vertices, N_{PV}. The lower range, $200\text{GeV} \leq p_{\text{T}}^{\text{jet}} < 300\text{GeV}$, contains hadronic boosted-objects, whereas the range $600\text{GeV} \leq p_{\text{T}}^{\text{jet}} < 800\text{GeV}$ is expected to contain top quarks, for which the decay products are merged within $R = 1.0$. Using the trimming technique, we can keep $\langle m^{\text{jet}} \rangle$ low even in the high pile-up condition.

4.5.2 Jet mass definitions

The jet mass is one of the most powerful discriminants for jet flavour tagging and is typically constructed with pure topo-cluster constituents (i.e. calorimeter information). The jet mass, m^{calo}, is defined as:

$$m^{\text{calo}} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2} \quad , \tag{4.15}$$

where E_i and $\vec{p}_i$ represent the energy and momentum of a topo-cluster i, respectively.

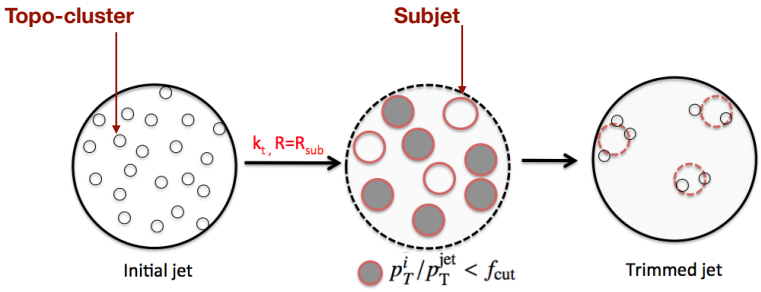
The angular spread in the decay products of a boosted massive particle scales as $1/p_{\text{T}}$. For highly-boosted W/Z bosons, this spread is comparable with the calorimeter granularity. To these highly-boosted W/Z bosons, we achieve better mass resolution using the track information as a result of the high spacial granularity of tracks reconstructed in the inner detector. We construct 'track-assisted' jet mass, m^{TA}, using the track information, defined as:

$$m^{\text{TA}} = m^{\text{track}} \times \frac{p_{\text{T}}^{\text{calo}}}{p_{\text{T}}^{\text{track}}} \quad , \tag{4.16}$$

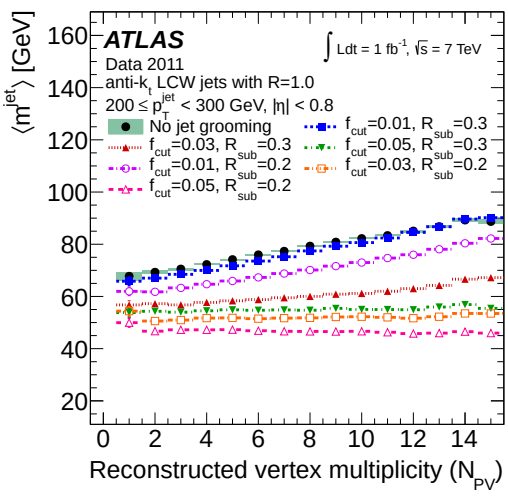
in which m^{track} and $p_{\text{T}}^{\text{track}}$ are the invariant mass and p_{T} calculated as the four-vector sum of the tracks associated with the large-R trimmed jet; $p_{\text{T}}^{\text{calo}}$ is the p_{T} of the original trimmed large-R jet, whose constituents are the topo-clusters.

To achieve the best resolution in overall p_{T} range, a new mass definition, m^{comb}, is defined using the linear combination of m^{calo} and m^{TA}.

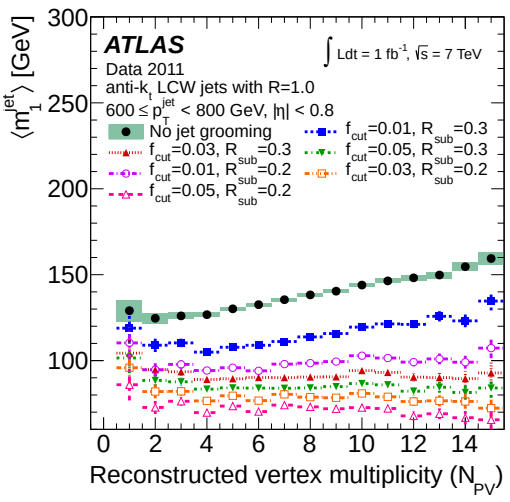
$$m^{\text{comb}} = m^{\text{calo}} \times w^{\text{calo}} + m^{\text{TA}} \times w^{\text{TA}}, \quad \left(\text{where } w_{\text{TA}} = \frac{\sigma_{\text{TA}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}}, w_{\text{calo}} = 1 - w_{\text{TA}} \right) \quad , \tag{4.17}$$



(a)



(b)



(c)

FIGURE 4.21: (a) Schematic view of the jet trimming procedure. Evolution of the mean trimmed and un-trimmed jet mass,$\langle m^{\text{jet}} \rangle$ in the central region $|\eta| < 0.8$ as a function of number of the reconstructed vertices: (b) for 200 GeV $\leq p_{\text{T}}^{\text{jet}} \leq$ 300 GeV, and (c) for 600 GeV $\leq p_{\text{T}}^{\text{jet}} \leq$ 800 GeV [128].

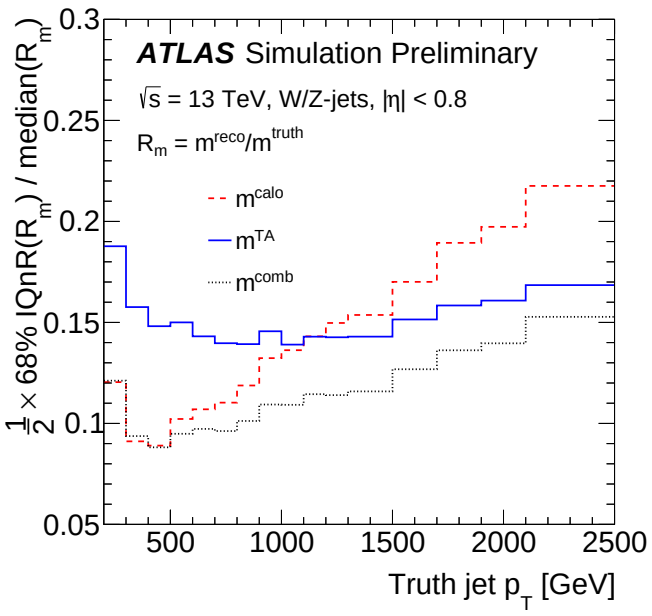


FIGURE 4.22: Jet mass resolution as a function of jet p_{T} for m^{calo}, m^{TA}, and m^{comb} for W/Z-jets [132]. The half of the 68% interquantile range (IQnR) divided by the median of the jet mass response is used as an outlier insensitive measure of the resolution. Due to the non-gaussian components in the jet mass distribution, the linear combination is not optimised in the lowest bin.

where σ_{TA} and σ_{calo} are the resolution of m^{TA} and m^{calo}, respectively. Figure 4.22 shows the mass resolution for $m^{\mathrm{calo}}, m^{\mathrm{TA}}$, and m^{comb}. In the low p_{T} region ($p_{\mathrm{T}} < 1.2$ TeV), the resolution of m^{calo} is better than that of m^{TA}. However, this resolution gets worse in higher p_{T} region because the granularity of the calorimeter is not enough. On the other hand, m^{TA} achieves better resolution than m^{calo} in high p_{T} region because m^{TA} is calculated using finer granularity inputs from the inner detector. Combining them, m^{comb} achieves the best resolution in overall p_{T} range. Thus, we adopt m^{comb} for the large-R jet mass in this analysis.

4.5.3 Jet substructure variable to identify boosted two prong jets

To distinguish the hadronic decays of boosted W/Z bosons from massive QCD jets, we adopt a jet substructure variable, D_2 in this analysis. A schematic view of 1-prong jets, which are dominantly from the QCD background and jets with a resolved 2-prong structure (subjets), such as from boosted W/Z decays, are shown in Figure 4.23. Two- and three-point energy correlation functions on a jet (e_2, e_3) enable us to resolve two hard subjets. Here are the definitions of the functions:

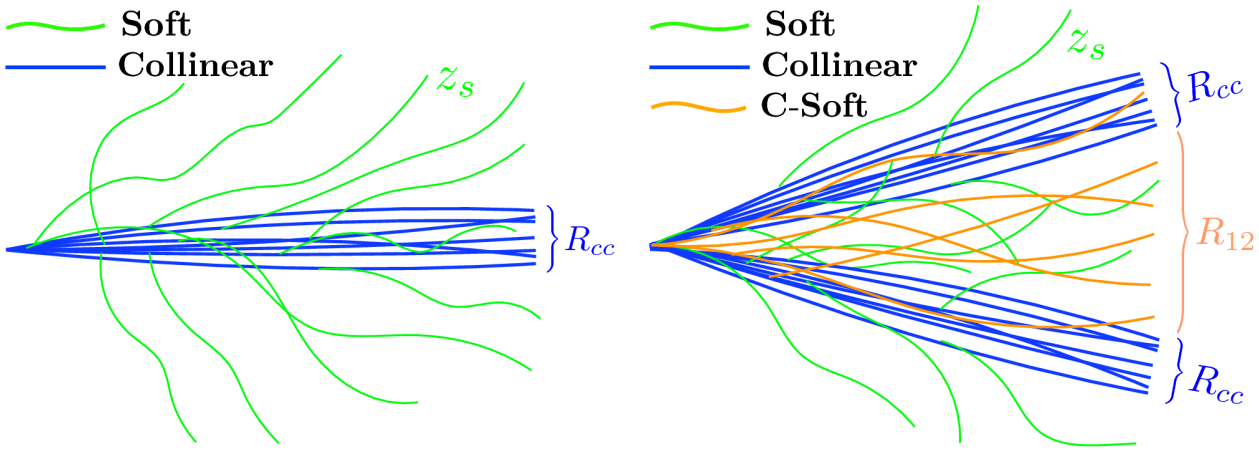
$$e_2^{\beta} = \frac{1}{p_{\mathrm{TJ}}^2} \sum_{1 \le i \le j \le n_{\mathrm{J}}} p_{\mathrm{T}i} p_{\mathrm{T}j} R_{ij}^{\beta} \quad , \tag{4.18}$$

$$e_3^{\beta} = \frac{1}{p_{\mathrm{TJ}}^3} \sum_{1 \le i \le j \le k \le n_{\mathrm{J}}} p_{\mathrm{T}i} p_{\mathrm{T}j} p_{\mathrm{T}k} R_{ij}^{\beta} R_{ik}^{\beta} R_{jk}^{\beta} \quad , \tag{4.19}$$

where p_{TJ} is the transverse momentum of the jet with respect to the beam, $p_{\mathrm{T}i}$ is the transverse momentum of particle i, and n_{J} is the number of particles in the jet. The boost-invariant angle $R_{ij}^2 = (\phi_i - \phi_j)^2 + (y_i - y_j)^2$ is the Euclidean distance in the azimuthal rapidity plane. For infrared and collinear (IRC) safety, the angular exponent β has to be greater than 0.

The 1- and 2-prong jets populate the (e_2, e_3) phase space regions where:

$$\begin{aligned} 1-\text{prong jet} &: (e_2)^3 \lesssim e_3 \lesssim (e_2)^2 \quad . \\ 2-\text{prong jet} &: 0 < e_3 \ll (e_2)^3 \quad . \end{aligned}$$



(a) (b)

FIGURE 4.23: (a) 1-prong jet, dominated by collinear (blue) and soft (green) radiation. The angular size of the collinear radiation is R_{cc} and the p_T fraction of the soft radiation is z_s. (b) 2-prong jet resolved into two subjets, dominated by collinear (blue), soft (green), and collinear-soft (orange) radiation emitted from the dipole formed by the two subjets. The subjets are separated by an angle R_{12} [133].

Background QCD jets dominantly populate the 1-prong region of the phase space, while the boosted W/Z signals dominantly populate the 2-prong region of phase space. The boundary between the background-rich and signal-rich regions is $e_3 \sim (e_2)^3$, thus we define a observable for discriminating boosted W/Z bosons from QCD jets as:

$$D_2^\beta \equiv \frac{e_3^\beta}{\left(e_2^\beta\right)^3} \quad . \tag{4.20}$$

Jets in the boosted W/Z signal events are characterised by a small value of D_2^β, while jets in the background events pre-dominantly have large D_2^β. In this analysis, we adopt $\beta = 1$ for the D_2 as a jet substructure variable to identify boosted two-prong jets.

4.5.4 Boson tagging

To identify the boosted W/Z-boson jets, a cut-based tagger is developed using two discriminants; m^{comb} and D_2, which represents '2-prongness'. The tagger provides a set of m^{comb} and D_2 selections that vary based on the p_T of the jet and provide an approximately constant signal efficiency, 50 and 80 %, over the p_T range.

For m^{comb}, a two-sided mass cut is used while a single-sided cut is applied for D_2. To provide a smooth cut function, optimised cut values as a function of jet p_T are provided using two different functions. A polynomial function is used to fit single-sided D_2 cut, and an empirical function, $\sqrt{(A/p_T + B)^2 + (C/p_T + D)^2}$, is used to fit the two-sided mass cut, where A, B, C and D are the fit parameters. The background rejection as a function of p_T of truth jet is shown in Figure 4.24. At $p_T = 1$ TeV, $\sim$ 70% of background jets are rejected by the optimised working point with 50% signal efficiency.

4.6 Missing transverse energy reconstruction

We use the missing transverse energy (E_T^{miss}) to reconstruct the transverse momentum carried by undetected particles; neutrinos for the analysis presented in this thesis. E_T^{miss} is reconstructed with two components; (1) hard objects, which are fully reconstructed and calibrated particles (electrons, photons, and muons, while hadronic-decaying τ leptons are included in jet term) and jets; (2) soft objects, which are reconstructed tracks associated

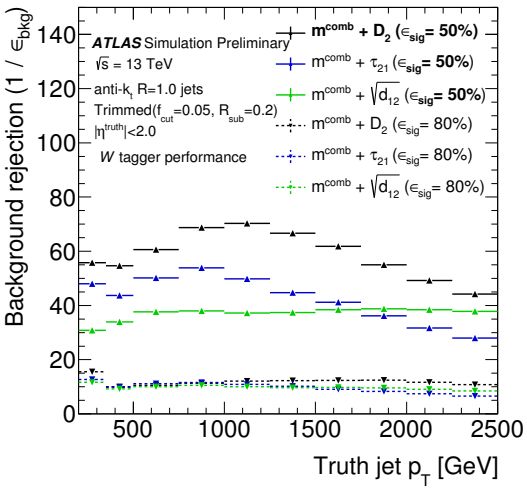


FIGURE 4.24: Background rejection in W-boson tagging as a function of p_{T} of truth jets [35]. Two working points are provided at 50 and 80 % signal efficiency.

with a primary vertex but not associated with the hard objects. Equation 4.21 describes each of the components;

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum\left(\vec{E}_{\mathrm{T}}^{jet} + \vec{E}_{\mathrm{T}}^{e} + \vec{E}_{\mathrm{T}}^{\mu} + \vec{E}_{\mathrm{T}}^{\gamma} + \vec{E}_{\mathrm{T}}^{soft}\right) \quad , \tag{4.21}$$

where $\vec{E}_{\mathrm{T}}^{jet}, \vec{E}_{\mathrm{T}}^{e}, \vec{E}_{\mathrm{T}}^{\mu}, \vec{E}_{\mathrm{T}}^{\gamma}$, and $\vec{E}_{\mathrm{T}}^{soft}$ represent the transverse energies of jets, electrons, muons, τ-leptons, photons, and soft components, respectively. Usually $E_{\mathrm{T}}^{\mathrm{miss}}$ denotes the norm of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, and we often decompose $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ to x and y directions and define the azimuth of the missing transverse energy as:

$$\tan\phi = \frac{E_y^{\mathrm{miss}}}{E_x^{miss}} \quad , \tag{4.22}$$

where E_x^{miss} and E_y^{miss} are the projection of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ on the x- and y-axis, respectively. Systematic uncertainties on each term in Equation 4.21 are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty. Figures 4.25 show the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions and resolution with and without genuine missing transverse momentum generated by undetectable particles.

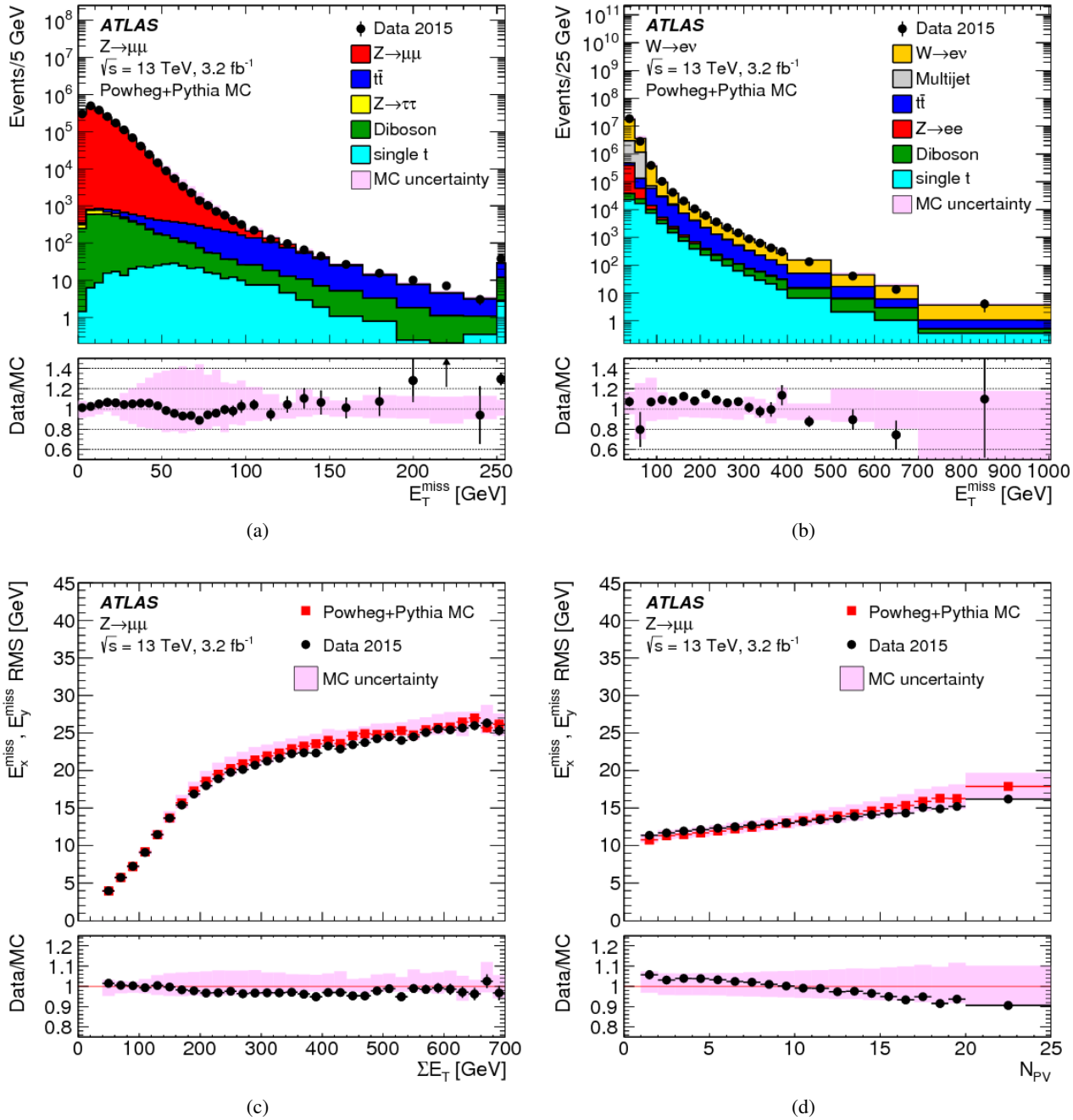


FIGURE 4.25: (a): Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in $Z \to \mu\mu$ events, (b): Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in $W \to e\nu$ events. The expectation from MC simulation is composed of all relevant backgrounds. The Data-to-MC ratio is shown in the bottom panels with the shaded areas indicating the total uncertainties for MC simulations. The RMS width of the $E_{x(y)}^{\mathrm{miss}}$ distributions (c) binned in $\sum E_{\mathrm{T}}$ and (d) binned in the number of primary vertices in $Z \to \mu\mu$ events.

Chapter 5

Jet energy calibrations

After the jet reconstruction, we apply several corrections to the reconstructed jets. The calibration procedure is divided into two large steps: MC-based calibrations and data-driven calibrations. We describe the calibration flow based on MC simulation in Section 5.1. From Section 5.2, the data-driven, called *in-situ*, calibrations are described.

5.1 Jet calibration flow and Monte-Carlo based calibrations

Figure 5.1 shows the flow of the ATLAS jet calibration scheme both for (a) small-R (anti-k_t $R = 0.4$) jets and (b) large-R jets. The calibration steps are largely divided into two stages: Monte-Carlo based and data-driven calibrations. We only apply two calibrations: the MC-based calibrations and the *in-situ* calibration, for large-R jets. Note that the trimming is applied before applying the jet energy calibrations. In this section, we review the MC-based calibrations: the origin correction, the pile-up corrections, the MC-based calibration, and the global sequential calibration.

5.1.1 Origin correction

We measure the energy of particles using the calorimeter, thus topo-clusters require to be assigned a direction to complete their four-vector. The default choice is to point them at the centre of the detector, a better assumption is, however, that they originated from the position of a primary vertex (Note that the length of beamspot along the beam line is ~50 mm). The origin correction accounts for this difference by finding the energy centre of a jet and then modify the jet so that the energy is unchanged but the direction originates from a primary vertex. This procedure results in an improvement in the η resolution of jets from 0.06 to 0.045 at $p_{\mathrm{T}} = 20$ GeV and from 0.03 to below 0.006 above 200 GeV [134].

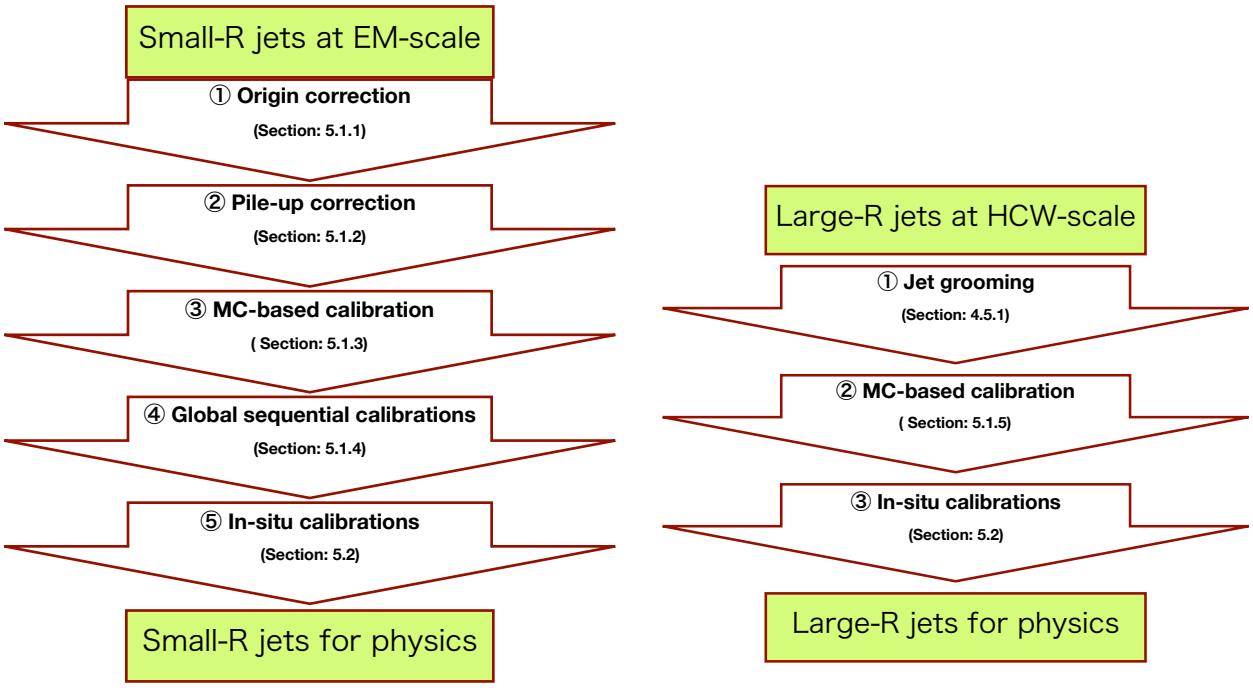
5.1.2 Jet area and residual pileup corrections

Following the origin correction, to reduce the effects of pileup on the jet calibration, we perform an area-based and residual subtraction methods. p_{T} corrected accounting for the pile-up effects, $p_{\mathrm{T}}^{\mathrm{corr}}$, is given by:

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{EM}} - \rho \times A - \alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \langle \mu \rangle \quad , \tag{5.1}$$

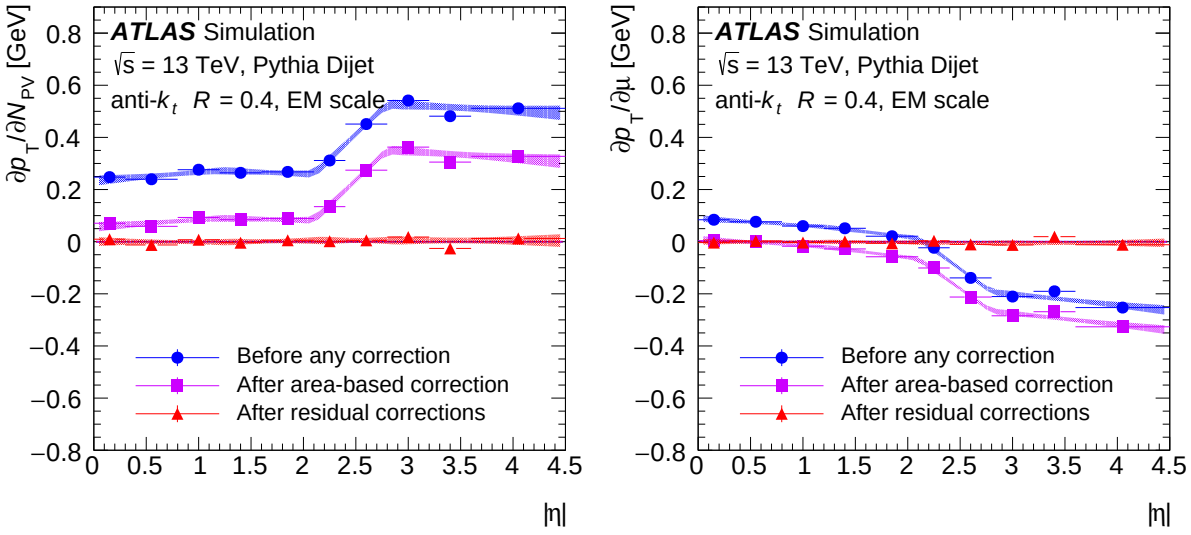
where $p_{\mathrm{T}}^{\mathrm{EM}}$ represents the un-calibrated jet p_{T} in the electromagnetic scale. The $\rho \times A$ term represents the area-based subtraction, where 'ρ' is the pile-up energy density and 'A' is the jet area in the (η, ϕ) plane.

After the area-based correction, there remains some small pileup dependence of the jet p_{T}, therefore an additional residual correction is required. The residual p_{T} dependencies on the number of primary vertices N_{PV} and on the average number of interactions per bunch crossing $\langle \mu \rangle$ are found to be fairly linear and independent on each other, thus the dependencies, represented by the term $(-\alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \langle \mu \rangle)$, are parametrized with N_{PV} and $\langle \mu \rangle$. The in-time pile-up dependence on N_{PV} is shown in Figure 5.2(a); the out-of-time pile-up dependence on $\langle \mu \rangle$ is shown in Figure 5.2(b). After the pile-up corrections, there is little pile-up dependence on jet p_{T}.



(a) (b)

FIGURE 5.1: Calibration flow for (a) anti-k_t $R = 0.4$ jets and (b) large-R jets. The calibration is applied to the four-momentum of the jet in each step except for the origin correction [134]. For the large-R jets, we only apply the absolute MC-based calibration before the *in-situ* calibrations [135].



(a) (b)

FIGURE 5.2: Dependence of anti-k_t $R = 0.4$ jet p_T on (a) the in-time pile-up (N_{PV} averaged over μ) and (b) out-of-time one. The dependence is shown binned in η before pile-up corrections (blue), after the area-based correction (violet) and after the the residual correction (red) [134].

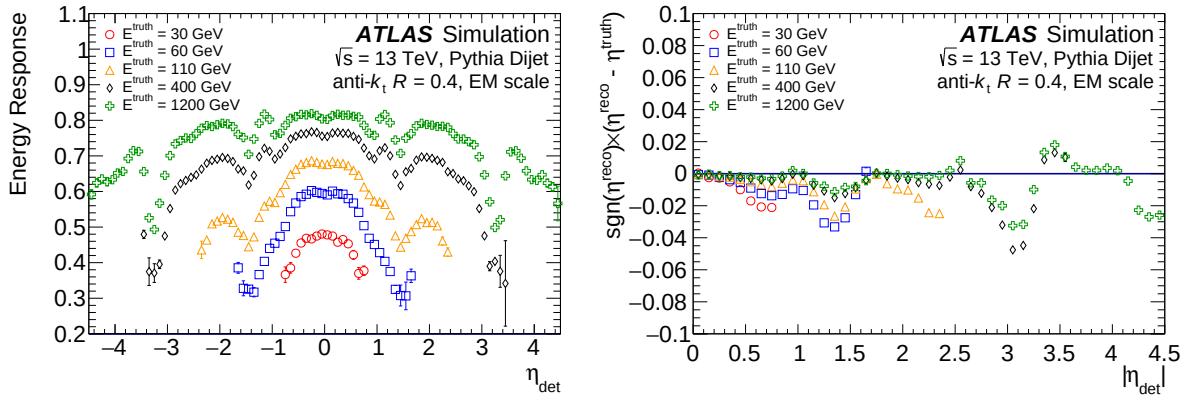
5.1.3 MC-based jet energy scale calibration

After the origin and the pile-up corrections are applied, the average energy response is calibrated to the truth jet energy by a MC-based jet energy calibration. The jet energy scale factors are derived from a MC simulation sample with inclusive isolated jets. The distribution of correction factors by the energy response = E^{reco}/E^{truth}, is shown as a function of η_{det}, which is defined as the jet η pointing from the geometric centre of the detector, in Figure 5.3(a). There are several gaps on the figure, which represent undetected particles in the transition region between the segments of the calorimeters. This calibration corrects the average of the reconstructed p_T of jets to their truth value.

Some bias are found in the reconstructed jet η as shown in Figure 5.3(b). The bias is largest in jets which traverse two calorimeter sub-sectors with different energy responses.The artefact increases the energy of one side of the jet with respect to the other, biasing the reconstructed four-momentum. The bias in the barrel-endcap ($|\eta_{det}| \sim 1.4$) and endcap-forward ($|\eta_{det}| \sim 3.1$) transition regions can be clearly seen in Figure 5.3(b) as a result of this effect. Thus a second correction needs to be applied to correct the difference between the reconstructed η_{reco} and truth η_{truth}, parametrized as a function of E_{truth} and η_{det}. The energy response is again used to derive the energy correction factors. Unlike the other calibration steps, the η calibration modifies only the jet p_T and η, not the full four-momentum.

5.1.4 Global sequential calibration (GSC)

Following the previous jet calibrations, it is observed that there is a difference between the closure of quark and gluon initiated jets; a quark-initiated jet often include hadrons with a higher fraction of the jet p_T that penetrate into the calorimeter, while a gluon-initiated jet typically contain more particles of softer p_T, leading to a lower calorimeter response and a wider transverse profile. To reduce the difference between the jet responses of quarks and gluons, further corrections, the global sequential calibrations (GSC) are applied. We apply the corrections depending on the topology of energy deposits in the calorimeter, tracking information, and muon spectrometer information. The GSC variables are listed in Table 5.1 and their distributions are shown in Figure 5.4. The corrections by the GSC variables are applied independently and sequentially without changing the average jet energy scale in the QCD inclusive jet sample used as a reference. The jet resolution, however, improves as more GSCs are applied. After the full GSCs are applied, the dependence of the jet response on each observable is reduced to less than 2%.



(a) (b)

FIGURE 5.3: (a) The average energy response as a function of η_{det} for jets of a truth energy of 30, 60, 110, 400 and 1200 GeV. The response is derived after the origin and pile-up corrections are applied. (b) The signed difference between the truth jet η^{truth} and the reconstructed jet η^{reco} due to the bias in the jet reconstruction. This bias is addressed with an η correction applied as a function of η_{det} and E_{truth} [134].

TABLE 5.1: Sequence of GS corrections used to improve the jet performance in each $|\eta|$ region.

	GSC$_{\text{calo}}$		GSC$_{\text{track}}$		GSC$_{\text{muon}}$
$\|\eta\|$ region	Correction 1	Correction 2	Correction 3	Correction 4	Correction 5
[0, 1.7]	f_{Tile0}	f_{LAr3}	n_{trk}	$width_{\text{trk}}$	N_{segments}
[1.7, 2.5]		f_{LAr3}	n_{trk}	$width_{\text{trk}}$	N_{segments}
[2.5, 2.7]		f_{LAr3}			N_{segments}
[2.7, 3.5]		f_{LAr3}			

Layer fractions $f_{\text{Layer}} = \frac{E_{\text{EM}}^{\text{Layer}}}{E_{\text{EM}}^{\text{jet}}}$

n_{trk} = number of associated tracks

$\text{width}_{\text{trk}}$: tracking jet width

$$\text{width}_{\text{trk}} = \frac{\sum_i p_{\text{T}}^i \Delta R(i, \text{jet})}{\sum_i p_{\text{T}}^i}$$

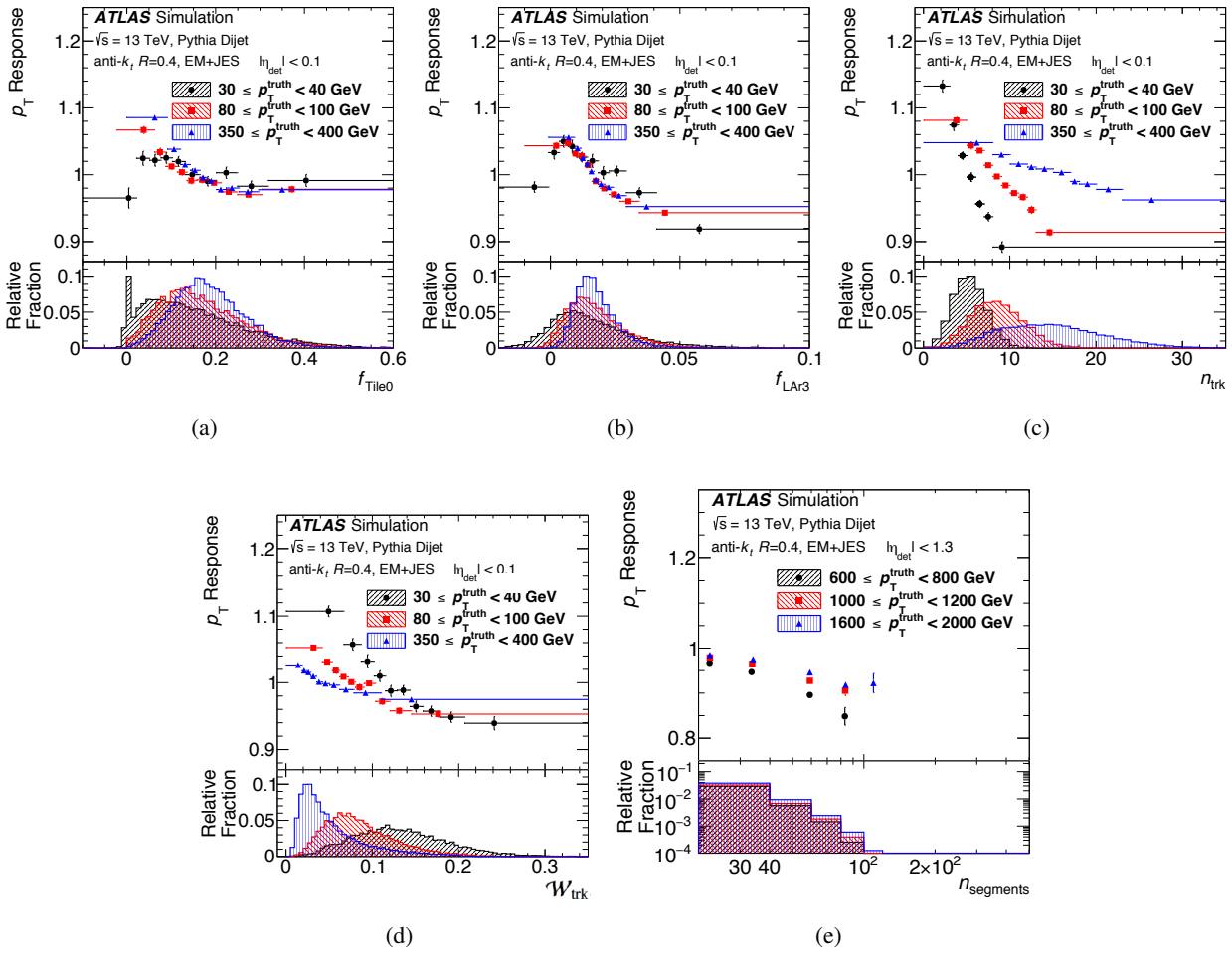


FIGURE 5.4: The average jet response in a simulation sample as a function of the GSC variables for three ranges of $p_{\mathrm{T}}^{\mathrm{truth}}$. (a) The fraction energy in the first Tile calorimeter layer , (b) the fractional energy in the third LAr calorimeter layer, (c) the number of tracks in a jet, (d) the p_{T}-weighted track width, and (e) the number of muon track segments in a jet [134]. The calorimeter distributions (a) and (b) are shown with no GSC corrections applied, the track-based distributions (c) and (d) are shown with the calorimeter-based corrections applied, and the muon spectrometer information (e) is shown with both the calorimeter-based and the track-based calibrations are applied.

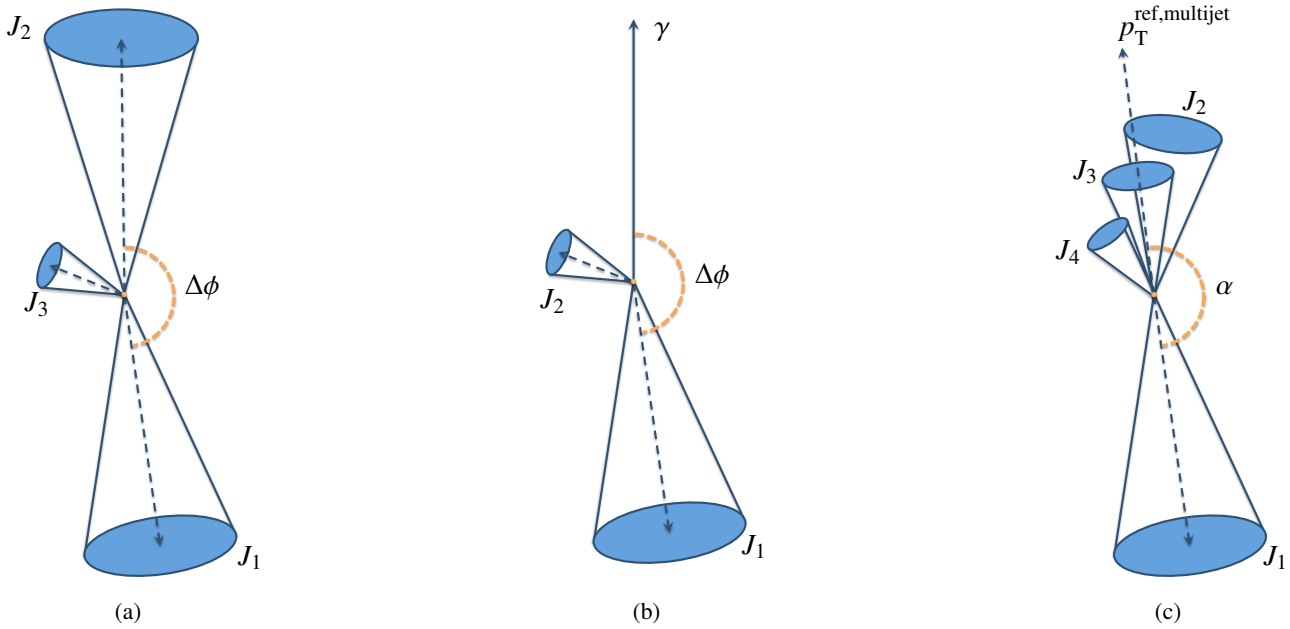


FIGURE 5.5: Schematic representation of the events used to constrain the jet energy scale (JES) and jet energy resolution: (a) a dijet event, (b) an ideal γ–jet event for use in the direct balance method, and (c) a multi-jet event with several jets recoiling against the leading jet [48]. In the dijet event, the third jet should have p_T of less than 50 GeV. We use the Z-jet balance method in the small-R jet energy calibration flow; the Z-jet event can be represented in a similar way with Figure (b).

5.1.5 MC-based calibrations for large-R jets

For the large-R jets, we only apply the MC-based JES corrections to restore the average reconstructed calorimeter jet energy scale and η to the truth jet energy scale and η. However, in addition to the JES, the jet invariant mass is also calibrated in the final step [128]. This is important when we use the jet mass in physics analyses because the jet mass is more sensitive to soft, wide-angle contributions, and to the merging and splitting. In this procedure, we measure the jet mass response after the standard MC-based jet energy scale calibration. The correction is derived using the same techniques as the jet energy correction. A correction parametrized in jet p_T, η, and mass is applied keeping the jet p_T constant.

5.2 *in-situ* jet energy calibrations

Following the MC-based calibrations, we apply a set of data-driven '*in-situ*' calibrations as the final stage of the jet energy calibration. In this section, we concentrate on the *in-situ* calibration methods for the large-R jets because the calibration performance of the large-R jets is more significant for the analysis than the performance of the usual anti-k_t $R = 0.4$ jets.

The strongest constraints on the p_T response are based on topologies where a jet recoils against a well-calibrated reference object or system, such as a photon and a small-R jet in a well-understood region (i.e. with lower p_T, and in the central part of the detector). The reference objects are selected depending on jet p_T to obtain a best performance. Transverse momentum conservation allows a determination of the jet energy resolution and the transfer of the transverse momentum scale of the reference object to the probe jet. The three topologies used in the calibrations are schematically depicted in Figure 5.5; (a) the dijet, (b) gamma-jet, and (c) multi-jet balance methods.

5.2.1 Determination of the jet energy resolution from dijet events

This subsection provides the methods used to measure the jet energy resolution (JER), with the topologies of Figure 5.5 (a), using the p_T balance between the leading and the sub-leading large-R jets. We define a p_T asymmetry

between the two jets as:

$$A = 2\frac{p_{\mathrm{T}}^{1} - p_{\mathrm{T}}^{2}}{p_{\mathrm{T}}^{1} + p_{\mathrm{T}}^{2}} \quad , \tag{5.2}$$

where p_{T}^{1} and p_{T}^{2} are the transverse momenta of the two leading jets in an event. The distribution of the asymmetry follows a gaussian function. The width of the distribution is related to the momentum resolution of each jet, and is estimated by an iterative fit with a gaussian function, changing the fitting range at every iteration:

1. the first fit is over the full histogram range (typically A=[-1,1]),
2. the second fit is performed in a window of 2 σ around the first fit result,
3. the third fit is performed in a window of 2 σ around the second fit result.

The width of the asymmetry distribution depends on multiple effects; (1) the detector resolution, (2) the intrinsic particle-level width, which arises due to energy flow not captured by the finite-size jets (out-of-cone radiation). Since the latter effect is uncorrelated with the detector response, the component of the asymmetry width due to the detector resolution can be determined by subtracting in quadrature the asymmetry width of particle-level ('truth-level') jets from that of reconstructed jets:

$$\sigma_{A,\mathrm{Det}} = \sqrt{\sigma_{A,\mathrm{reco}}^{2} - \sigma_{A,\mathrm{truth}}^{2}} \quad . \tag{5.3}$$

If the two jets are both central ($|\eta| < 0.8$), they have approximately the same resolution. In this case, both jets contribute the same amount to the asymmetry distribution, the relative jet-p_{T} resolution is defined by:

$$\frac{\sigma(p_{\mathrm{T}}^{\mathrm{avg}})}{p_{\mathrm{T}}^{\mathrm{avg}}} \sim \frac{\sigma(A)}{\sqrt{2}} \quad . \tag{5.4}$$

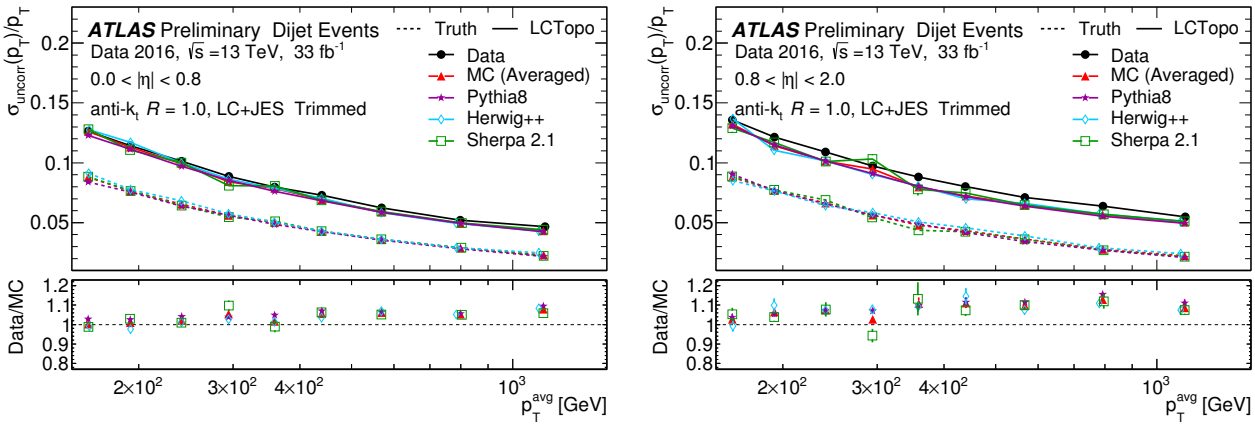
The resolution of the forward jets (in the region $0.8 < |\eta| < 2.0$), which is significantly different from that of central one, is extracted from the width of the asymmetry distribution in events where a central jet balances with a forward jet. The result for the resolution of the central jets is corrected by subtracting the asymmetry of the central dijet systems:

$$\frac{\sigma(p_{\mathrm{T}}^{\mathrm{fwd}})}{p_{\mathrm{T}}^{\mathrm{avg}}} = \sqrt{\sigma(A^{\mathrm{ctl,fwd}})^{2} - \frac{\sigma(A^{\mathrm{ctl,ctl}})^{2}}{2}} \quad , \tag{5.5}$$

where the labels (ctl) and (fwd) refer to the central and forward detector. The measured resolution of the truth jets is indicated in Figure 5.6. An example of the fit on 2016 data is shown as the solid black curve with the round markers, where the measured relative resolution $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ is plotted as a function of the average p_{T} of the two jets, $p_{\mathrm{T}}^{\mathrm{avg}}$. The other curves show the result of simulations based on three MC generators: Pythia8, Herwig++, and Sherpa 2.1. The ratio of the jet energy resolution between data and MC simulation samples, is shown Figure5.6 as well. For each of the generators, the width of the detector-level asymmetry is shown as a solid line, while the particle-level asymmetry is indicated by a dashed line.

The online selection of the events relies on a combination of jet triggers. This selection includes triggers based on small-R jets and large-R jets, with p_{T} thresholds ranging from 110 GeV to 420 GeV. The triggers are chosen to be practically 100% efficient in the p_{T} range considered. Dijet events with two large-R jets in opposite side of the detector in the transverse direction are selected by requiring $\Delta\phi > 2.5$ to ensure that they are back-to-back. The third large-R jet in the event should have p_{T} of less than 50 GeV. These requirements ensure a proper dijet topology, where the p_{T} of both jets balance with each other.

The relative resolution estimated from the p_{T} asymmetry of the two jets, after the subtraction of the truth resolution, is shown in Figure 5.7. Figure 5.7(a) corresponds to the central detector region ($|\eta|$ <0.8), and 5.7(b) corresponds to the the pseudo-rapidity interval between 0.8 and 2.0. The relative resolution $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ is extracted from the asymmetry distribution. The total uncertainty on the measurement is shown as the green band. The relative JER is estimated by comparing the MC expectations to the resolution measured in data. The central position of the line is computed from an average over all MC generators, while the envelope of the MC predictions is indicated with the red band. The result in the central region is in reasonable agreement with the prediction. The residual uncertainty on the jet energy resolution is approximately 2% at 300 GeV. The uncertainty increases slowly at higher p_{T}, but remains well below 5% up to 1 TeV. The resolution of the forward jets is smaller in the MC simulation



(a) (b)

FIGURE 5.6: Comparison between jet p_T resolution in data and in simulation, uncorrected for truth-level intrinsic asymmetry resolution, as a function of the average jet p_T (black, round markers). The measurement is compared to predictions from Monte Carlo simulation based on three generators: PYTHIA8 in light blue, HERWIG++ in green and SHERPA 2.1 in purple. Also an unweighted average of the three is shown. Statistical errors are usually smaller than the size of the marker. The resolution at the truth-level is also shown, with dashed lines with the same colour coding. The plot (a) corresponds to the central region ($|\eta| < 0.8$), and (b) corresponds to the forward region ($0.8 < |\eta| < 2.0$) [48]

samples, typically by 5-10%. The choice of the MC generator has a small effect on the resolution. The deviation from unity is never larger than 10-15%.

The total uncertainty on the *in-situ* determination of the JER is shown in Figure 5.8 as a function of the average p_T in two pseudo-rapidity regions. A breakdown of the uncertainties into individual sources is presented as well. The p_T scale of the small-R jets is varied within its uncertainty, leading to a 2-3% variation in the measured resolution, labelled as "JES uncertainty" in Figure 5.8. The $\Delta\phi$ requirement is also varied from 2 to 3, which has an effect primarily at low p_T. Finally, the MC modelling uncertainty is estimated as the variation in the result for different generators (PYTHIA8 is chosen as a nominal sample, HERWIG++ and textscSherpa 2.1 as the variations).

5.2.2 Photon-jet balance calibration

The large-R jet energy scale can be constrained using the ideal photon–jet (γ–jet) final state as shown in Figure 5.5(b), because the energy of the photons can be measured more precisely in the electromagnetic calorimeter than the jets. In Run-1 at $\sqrt{s} = 7$ and 8 TeV, the photon energy calibration had an uncertainty of $\sim 0.3\%$ [136]. The photon calibration in Run-2 used in this thesis has similar uncertainties [118].

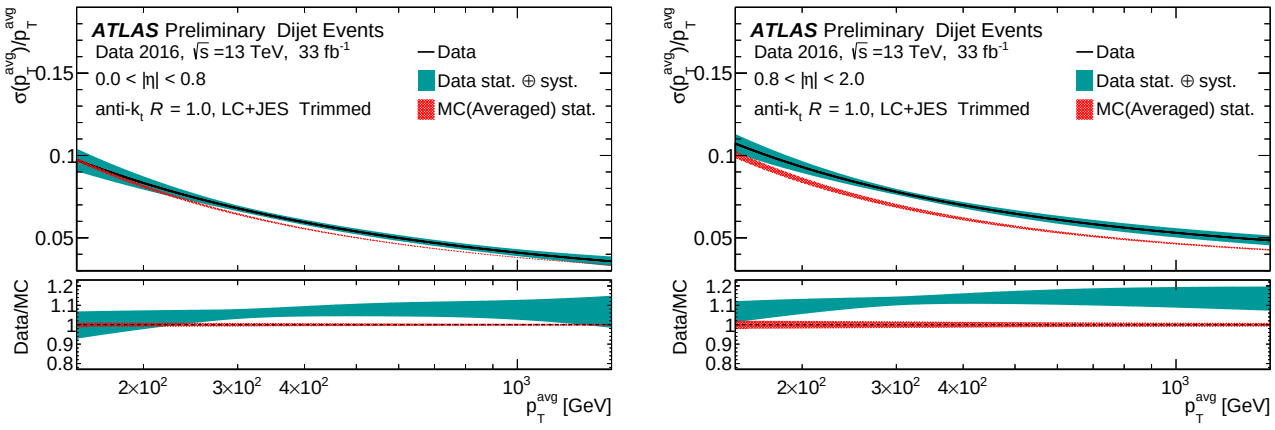
The γ–jet method [137] is based on the balance between the large-R jets and the photons:

$$B = \frac{p_T^{\text{jet}}}{p_T^{\text{ref}}} \quad , \tag{5.6}$$

where the reference momentum $p_T^{\text{ref}} = p_T^{\gamma}\left|\cos(\Delta\phi)\right|$ is a component of p_T^{γ} collinear to the jet. The balance B is calculated in both collision data and MC simulation samples. The double ratio of data and MC is defined as follows:

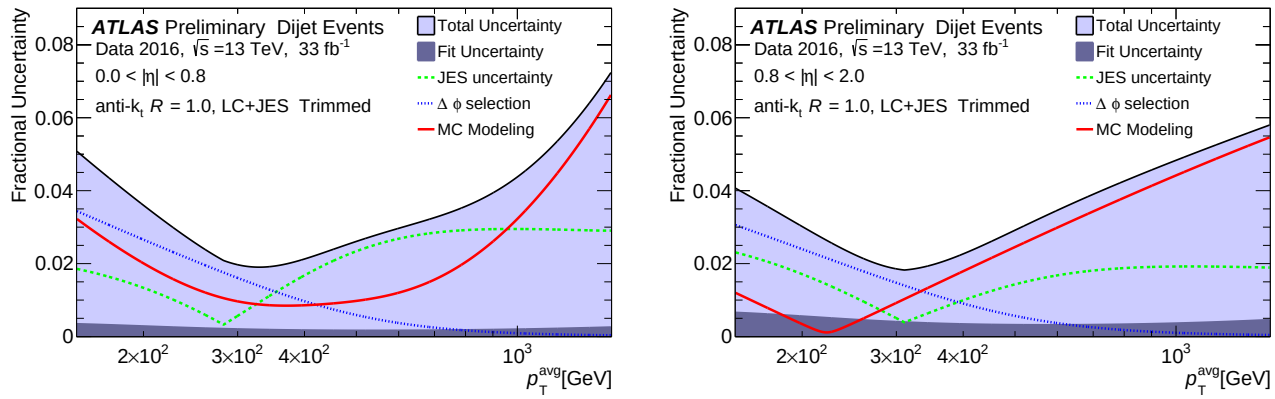
$$\frac{\langle B_{\text{data}}\rangle}{\langle B_{MC}\rangle} \equiv \frac{\langle p_T^{\text{jet}}/p_T^{\text{ref}}\rangle_{\text{data}}}{\langle p_T^{\text{jet}}/p_T^{\text{ref}}\rangle_{\text{MC}}} \quad , \tag{5.7}$$

which defines a residual mis-modelling in the jet energy scale calibration. If the reference photon is well measured experimentally and correctly modelled in simulation, any deviation in the double ratio can be attributed to the mis-modelling of the jet response in the MC simulation.



(a) (b)

FIGURE 5.7: Comparison of the measured jet p_{T} resolution in data and Monte Carlo simulation, as a function of the average jet p_{T}, after subtraction of the truth resolution. Figure (a) corresponds to the central region ($|\eta| < 0.8$), and (b) to the pseudo-rapidity interval of $0.8 < |\eta| < 2.0$. The green error band represents the statistical and systematic uncertainties added in quadrature. The red band shows the average of three Monte Carlo simulations and their envelope. The smaller panels below in each plot show t the ratio of the resolution in data the Monte Carlo simulation [48].



(a) (b)

FIGURE 5.8: The uncertainty on the ratio of the jet p_{T} resolution measured in dijet events and in Monte Carlo simulation as a function of the average jet p_{T}. The fit uncertainty is indicated with the dark filled region, the total uncertainty with the light colour. Contributions from three sources are estimated separately by propagating uncertainty on the p_{T} scale to the measurement, by varying the $\Delta\phi$ selection and by changing the Monte Carlo generator. Figure (a) corresponds to the central region ($|\eta| < 0.8$), and (b) to the forward region ($0.8 < |\eta| < 2.0$) [48].

TABLE 5.2: Summary of the event selections for the *in-situ* transverse momentum constraint in γ–jet events. Here J_1 refer to the leading large-R jet, and J_2 to the leading small-R jet. $E_{\mathrm{T}}^{\mathrm{iso}}$ is the calorimeter energy in a 0.4 cone around the photon, excluding a smaller region around the photon [48].

Object	Selection	Description
Photon (γ)	Single photon triggers	Trigger
	$p_{\mathrm{T}}^{\gamma} > 25$ GeV, $\lvert\eta^{\gamma}\rvert < 1.37$	Pre–selection
	`Tight` photon ID	Identification
	$E_{\mathrm{T}}^{\mathrm{iso}} < 0.022\, E_{\mathrm{T}} + 2.45$ GeV	Isolation
Large-R Jets (J_1)	$p_{\mathrm{T}}^{J_1} > 15$ GeV	Pre–selection
	$\Delta R\,(\gamma,\ J_1) > 0.2$	γ Overlap Removal
Small-R Jets (J_2)	`Loose` quality requirements	Jet Cleaning
	JVT > 0.59 for $p_{\mathrm{T}} < 60$ GeV, $\lvert\eta\rvert < 2.4$	Pileup Removal
	$\Delta R\,(\gamma,\ J_2) > 0.4$, $\Delta R\,(J_1,\ J_2) > 1.4$	Overlap Removal
Muons (μ)	$\Delta R\,(J_1,\ \mu) > 1.0$	J_1 Overlap Removal
Topological	$\Delta\phi(\gamma,\ J_1) > 2.8$	Separation Angle
	$p_{\mathrm{T}}^{J_2+\mu} < \max\left(15\text{ GeV},\ 0.1\ p_{\mathrm{T}}^{\mathrm{ref}}\right)$	Extra–radiation Veto

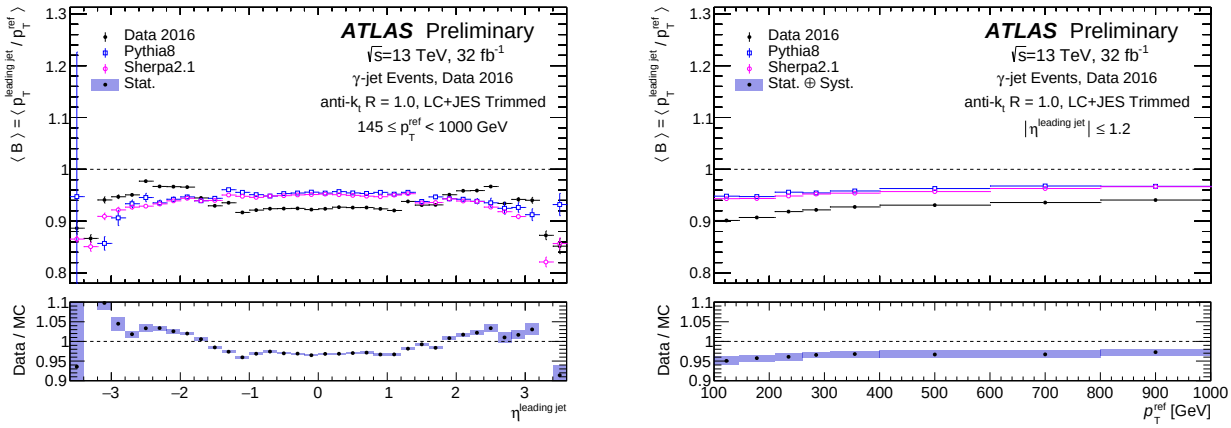
The γ-jet measurement follows the measurement in Reference [137], with some modifications and updates to match the latest configurations. Events are selected by requiring the presence of a photon satisfying the tight identification requirements [138] (In the tight identification requirement, such as the shower shapes in the first layer of the calorimeter are examined.) and a large-R jet that satisfies the requirements summarised in Table 5.2. For the online event selection, a combination of single photon triggers with p_{T} thresholds ranging from 20 GeV to 140 GeV is used. These cuts, in particular that on the azimuthal angle between the probe jet and the photon, are designed to select events where the transverse momentum is predominantly carried by the leading large-R jet, J_1, and the leading photon, γ. We apply a photon 'purity' correction to the mean of the balance in results in data to correct for contamination by jets or electrons that may skew the nominal p_{T} balance. The photon purity correction is derived using the 'side-bands' of the photon isolation[1] and ID[2], known as the 'ABCD' method [139].

The γ–jet balance method provides a precise constraint on the average of the jet p_{T} scale. The mean value of the balance B (defined in Equation 5.6) between the large-R jet and the recoiled photon is extracted from data and MC simulation samples by fitting the distribution of the balance with a gaussian function in bins of reference p_{T} and η. The uniformity of the large-R jet response across the detector geometry is studied; the γ–jet balance results as a function of the pseudo-rapidity of the probe jet is shown in Figure 5.9(a). The result in the upper panel shows the values of B for the γ–jet balance distribution in data and MC simulation samples; the smaller panel below shows the ratio of data and MC. The relative response is flat at the percent level in the central region to $\lvert\eta\rvert < 1.2$. The measured response in the forward region, however, shows a considerable structure. The MC simulation samples fail to correctly reproduce this structure, leading to a gradual increase of the data/MC ratio between $\lvert\eta\rvert \sim 1.2$ and ~ 2.5. In Figure 5.9(b) the result is plotted as a function of the reference p_{T}. We use the ratio between the result obtained in data and MC simulation to correct the jet energy scale.

The systematic uncertainties on the γ–jet *in-situ* p_{T} scale arise in four main categories; extra radiation outside the jet, the measurement of the photon, the presence of the pileup jets, and the MC physics modelling. The effects from the extra radiation on the balance are assessed by varying the topological selections and the overlap removal $\Delta R\,(J_1,\ j_2)$ between the large and small-R jets. The effects of the photon measurement are assessed by varying the energy scale and resolution of the photon calibration, as well as by varying the measured photon purity. The effects from the pileup jets on the calibration are estimated by varying the threshold for the jet–vertex–tagger (JVT, described in subsection 4.4.3) [125] selection threshold on the small-R jets. Lastly, the analysis is repeated with alternative SHERPA2.1 MC samples to assess the modelling uncertainty on the nominal PYTHIA8 samples. As shown

[1]The calorimeter isolation variable $E_{\mathrm{T}}^{\mathrm{iso}}$ is defined as the sum of the E_{T} of topological clusters deposited in a cone of size $\Delta R = 0.4$ around the photon candidate, excluding an area of $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ centred on the photon cluster, and subtracting the expected photon energy deposit outside of the excluded area.

[2]The photon identification decision is based on a set of shower shape variables computed from energy deposits in the first and second layers of the electromagnetic calorimeter, and from leakage to the hadronic calorimeter.



(a) (b)

FIGURE 5.9: The ratio of the mean values of the γ–jet balance distribution in data and Monte Carlo simulation (a) versus η of the large-R jet, and (b) versus reference p_{T} (the photon p_{T} scaled by the cosine of the angle between the photon and the jet). The blue uncertainty band on the relative response includes the statistical and systematic uncertainties [48].

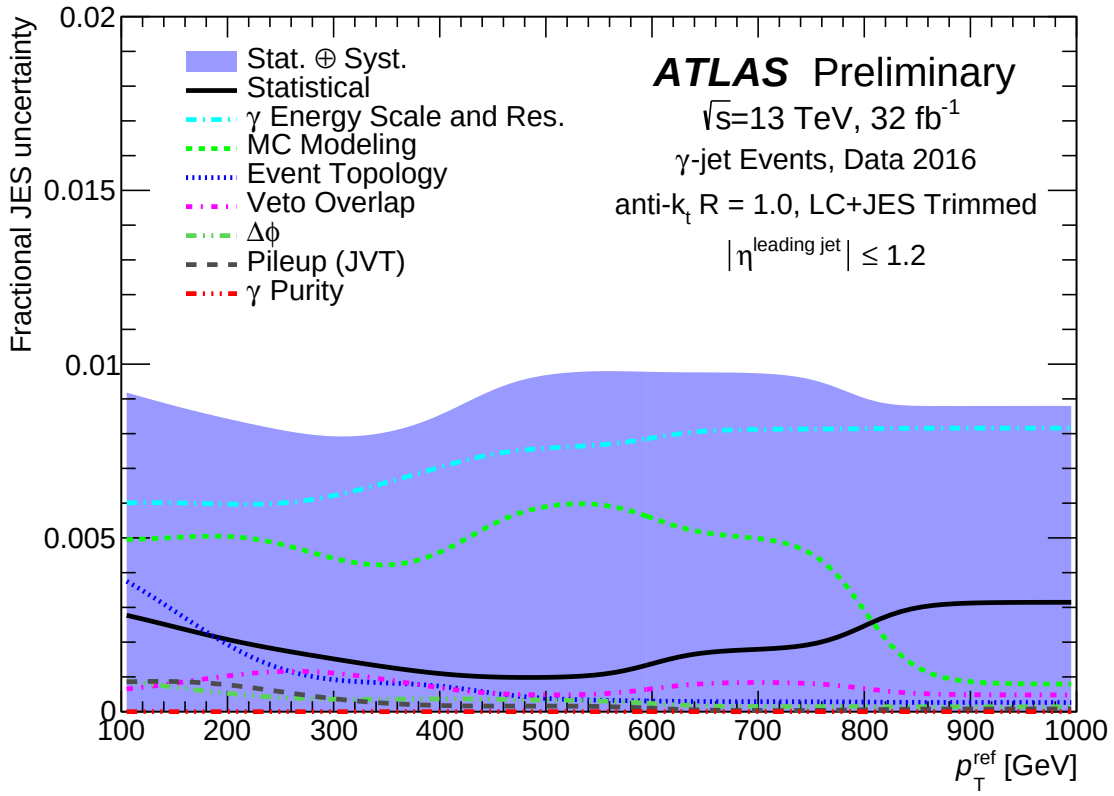
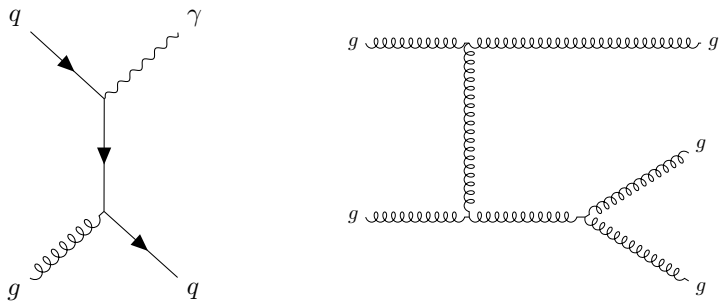


FIGURE 5.10: A break-down of the systematic uncertainties on the *in-situ* constraint on the jet energy scale obtained with the γ–jet method, as a function of the reference p_{T} [48].



(a) (b)

FIGURE 5.11: Typical diagrams of (a) the γ–jet and (b) the multi-jet events.

in Figure 5.10, the overall combined systematic and statistical uncertainty is approximately 1% for the $p_{\mathrm{T}}^{\mathrm{ref}}$ range from 100 GeV to 1 TeV. The photon energy scale uncertainty is the dominant source over most of the p_{T} range. This uncertainty is of the same order as the total JES uncertainty for $R = 0.4$ jets in Run 1 [137, 140].

5.2.3 Multi-jet balance calibration

To search for the diboson resonances in the higher mass range (resonance mass > 2 TeV), it is essential to determine p_{T} of highly-boosted W and Z bosons ($p_{\mathrm{T}} > 1$ TeV) precisely. As photons are produced via the electromagnetic interaction, as shown in Figure 5.11(a), the γ–jet events are statistically limited in the higher p_{T} range ($p_{\mathrm{T}} > 1$TeV). To extend the p_{T} range of the jet energy calibration to higher p_{T}, multi-jet events are used: the reference p_{T} is the vector sum of the $p_{\mathrm{T}}^{\mathrm{recoil}}$ of multiple, fully-calibrated jets balancing with a single, energetic large-R jet [140]. A typical example of the multi-jet events is shown in Figure 5.11(b), which is produced only via the strong interaction. A schematic representation of the event topology used in this method is shown in Figure 5.5(c) and Figure 5.12.

There are two options:

- to use the small-R jets as the recoil jets,
- to use the large-R jets as the recoil jets.

p_{T} of the small-R jets are well-calibrated up to 2.0 TeV; the small-R jets are already fully calibrated including the small-R multi-jet balance method. We decided to use the small-R jets as the recoil jets, thereby we can calibrate the energy of the large-R jets as high as possible and minimise the large-R JES uncertainties.

The transverse momentum balance, B_{MJB} is defined as:

$$B_{\mathrm{MJB}} = \frac{p_{\mathrm{T}}^{\mathrm{leading}}}{p_{\mathrm{T}}^{\mathrm{recoil}}} \quad , \tag{5.8}$$

where $p_{\mathrm{T}}^{\mathrm{leading}}$ is the transverse momentum of the leading large-R jet and $p_{\mathrm{T}}^{\mathrm{recoil}}$ is the vectorial sum of the transverse momenta of the 'recoil system' with small-R jets. The mean value of the B_{MJB} is measured both in data and MC simulation samples in bins of $p_{\mathrm{T}}^{\mathrm{recoil}}$. The data-to-MC double ratio $B_{\mathrm{MJB}}^{\mathrm{data}}/B_{\mathrm{MJB}}^{\mathrm{MC}}$ allows the estimation of the response of high-p_{T} jets.

Events used in the multi-jet balance (MJB) analysis are selected online by triggering on the leading small-R jet, using the single jet triggers. For each $p_{\mathrm{T}}^{\mathrm{recoil}}$ bin, a small-R jet trigger with the optimal threshold of jet p_{T} is used to select events efficiently. The triggers used for 200 GeV < $p_{\mathrm{T}}^{\mathrm{recoil}}$ < 550 GeV are prescaled, whereas an unprescaled jet trigger with p_{T} threshold of 380 GeV is used for $p_{\mathrm{T}}^{\mathrm{recoil}}$ > 550 GeV. The other selections are summarised in Table 5.3. For small-R jets with $p_{\mathrm{T}} < 60$ GeV, the JVT [125] selection is applied to suppress the pileup jets. The large-R probe jet is required to have $|\eta_{\mathrm{det}}| < 1.2$, while the small-$R$ jets that constitute the recoil system are required to have $|\eta_{\mathrm{det}}| < 2.8$. To select events with the multi-jet recoil systems, the leading jet in the recoil system ($p_{\mathrm{T}}^{\mathrm{jet2}}$) is allowed to have less than 80 % of the total transverse momentum of the recoil system (asymm$^{p_{\mathrm{T}}} = p_{\mathrm{T}}^{\mathrm{jet2}}/p_{\mathrm{T}}^{\mathrm{recoil}} <$ 0.8). Furthermore, the angle in the azimuthal plane between the leading large-R jet and the vector defining the

TABLE 5.3: Summary of the event selection for the multi-jet balance analysis [48]. The variables used in the selections are defined in Figure 5.12.

Variable	Selection
Pseudo-rapidity	$\|\eta\| < 1.2$ (probe jet), $\|\eta\| < 2.8$ (recoil jets)
Separation angle α	$\|\alpha - \pi\| < 0.3$
Separation angle β	$\Delta R > 1.5$
Number of recoil jets	≥ 2
Asymmetry recoil system	asymm$^{p_T} < 0.8$

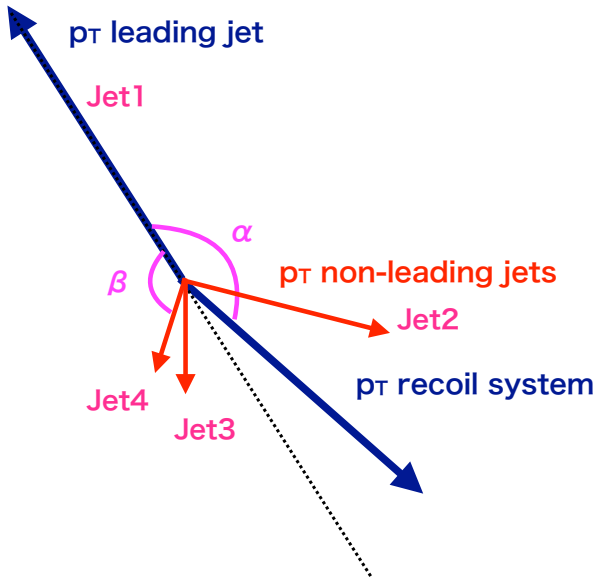


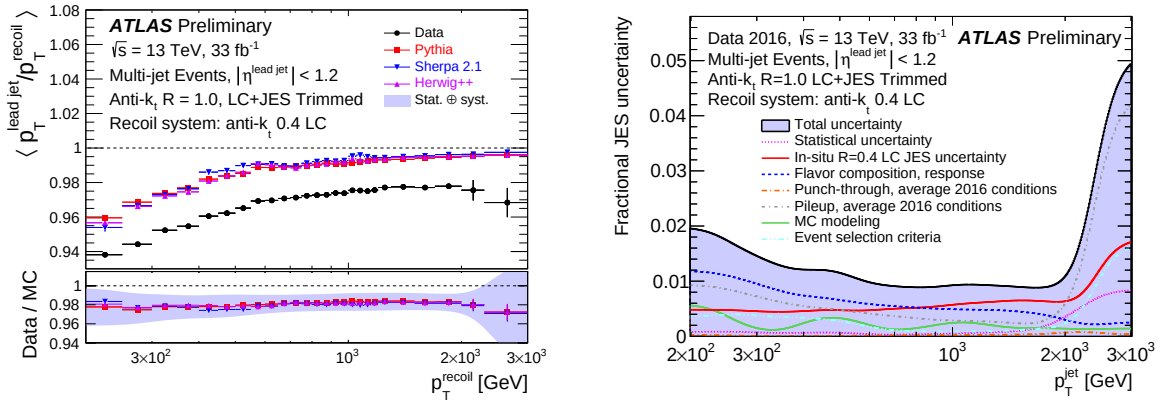
FIGURE 5.12: Schematic view of the variables used in the multi-jet balance selections.

recoil system, α, is required to satisfy $|\alpha - \pi| < 0.3$, and β, the distance in η-ϕ between the leading large-R jet and the nearest small-R jet from the recoil system, is required to be greater than 1.5.

Figure 5.13(a) shows the distribution of the mean transverse momentum balance for the leading-p_T large-R jets. The balance increases from approximately 0.94 at p_T of 200 GeV to 0.98 for jets with p_T between 1-3 TeV. The MC simulation samples show a similar trend, but have a less pronounced slope. As a result, the data/MC ratio retains a slight slope.

The shaded band in the data/MC ratio plot of Figure 5.13(a) represents the total uncertainty of the measurement. The reference system in the multi-jet-balance method is made of small-R jets. They are calibrated using several *in-situ* approaches described in Reference [137, 140]. Thus, the JES uncertainty of the recoil system depends on the systematic and statistical uncertainties on each *in-situ* procedure. To propagate the uncertainties to the multi-jet-balance method, all input components are individually varied by $\pm 1\sigma$ and a full iterative analysis procedure is taken. In this study, no assumption is posed on the flavour of the recoil jets (originating from a gluon, a light quark, or a heavy-flavour quark); this is also a source of the systematic uncertainty. The uncertainty on the multi-jet-balance observable due to the jet flavour response is evaluated by propagating the small-R jet flavour response uncertainties.

The event selection criteria and the modelling in the event generators directly affect the p_T balance used to obtain the multi-jet-balance results. The impact of the event selection criteria is estimated by shifting each cut point by a specified amount in both direction and observing the change in the results. Using a similar approach to the systematic uncertainties in the small-R *in-situ* analysis, the p_T threshold for recoil-jets is shifted by ± 5 GeV, the asymmetry of p_T asymm$^{p_T} = p_T^{jet2}/p_T^{recoil}$ is shifted by ± 0.1, the angle α is shifted by ± 0.1, and the angle β is shifted by ± 0.4. The uncertainty due to the MC modelling of the multi-jet events is estimated from a symmetrised envelope of multi-jet-balance corrections obtained by comparing the nominal results obtained with PYTHIA8 to those obtained with SHERPA2.1 and HERWIG++.



(a) (b)

FIGURE 5.13: (a) Mean transverse momentum balance for leading-p_{T} large-R jets ($|\eta| \leq 1.2$) balanced against a system with at least two small-R jets ($p_{\mathrm{T}} \geq 25$ GeV, $|\eta| \leq 2.8$), as a function of the p_{T} of the recoil system. In the main panel the measured balance (black, closed circles) is compared to the prediction in Monte Carlo simulations based on several generators: PYTHIA8 (red, square markers), SHERPA2.1 (blue, triangular markers) and HERWIG++ (pink, triangular markers). In the smaller panel the ratio of data and MC is presented. The shaded band indicates the total uncertainty on the measurement. (b) The fractional uncertainty on the jet energy scale derived from leading-p_{T} large-R jets with $p_{\mathrm{T}} \geq 300$ GeVand $|\eta| \leq 1.2$ balanced against a system with at least two small-R jets ($p_{\mathrm{T}} \geq 25$ GeV, $|\eta| \leq 2.8$), as a function of the p_{T} of the recoil system [48].

Figure 5.13(b) shows the distribution of the fractional uncertainties on the jet energy scale derived for the leading-p_{T} large-R jets recoiling against multiple small-R jets. The total fractional uncertainty is comprised of several sources: the uncertainty on the p_{T} scale of the jets in the recoil system; the uncertainty on the flavour composition of the calibration samples and response uncertainties (combined); the uncertainties on the schemes used to match pileup condition in data and MC; the uncertainties on the 'punch-through' jets, which are hadronic or electromagnetic showers not contained in the calorimeter volume, which results in tracks in the muon detectors; the modelling uncertainty accessed as the differences between PYTHIA8 and HERWIG++, SHERPA2.1; and finally the uncertainties on the event selection criteria, accessed by varying the cut values.

The results presented in this subsection are the first results of the large-R MJB. By these results, we calibrate the W and Z bosons decaying hadronically up to 3 TeV within the systematic uncertainty of 5-6 %.

5.3 Combination procedure

The measurements of the relative response of the jet p_{T} obtained with different *in-situ* methods are combined to constrain the JES uncertainties over a broad range, following the combination procedure described in more detail in Reference [141]. The data-to-MC p_{T} response ratios obtained from the γ–jet and multi-jet balance are combined to produce a single p_{T} response measurement. The systematic uncertainties of the measurements are propagated through the combination procedure to yield an associated jet p_{T} scale uncertainty.

The *in-situ* p_{T} measurements are given in bins of transverse momentum of $p_{\mathrm{T}}^{\mathrm{ref}}$. The combination proceeds in three steps which take into account correlations between uncertainties and inconsistencies between the *in-situ* methods:

1. **Toy Monte Carlo method:** To combined the γ–jet and multi-jet balance measurements, we use pseudo experiments from the nominal Monte Carlo sample carried out to represent the combination of the measurements. These pseudo-experiments are used to consistently propagate all uncertainties to the combined results taking into account the correlation between the measurements. The pseudo-experiments represent the full list of available measurements and contain the full chain of the uncertainty evaluation including interpolation and averaging (described in the following steps).

2. **Interpolation:** The relative p_{T} response (i.e. the ratio of measured and simulated response) is defined in fine $p_{\mathrm{T}}^{\mathrm{ref}}$ bins for each *in-situ* method, and interpolated with splines based on first- or second-order polynomials.

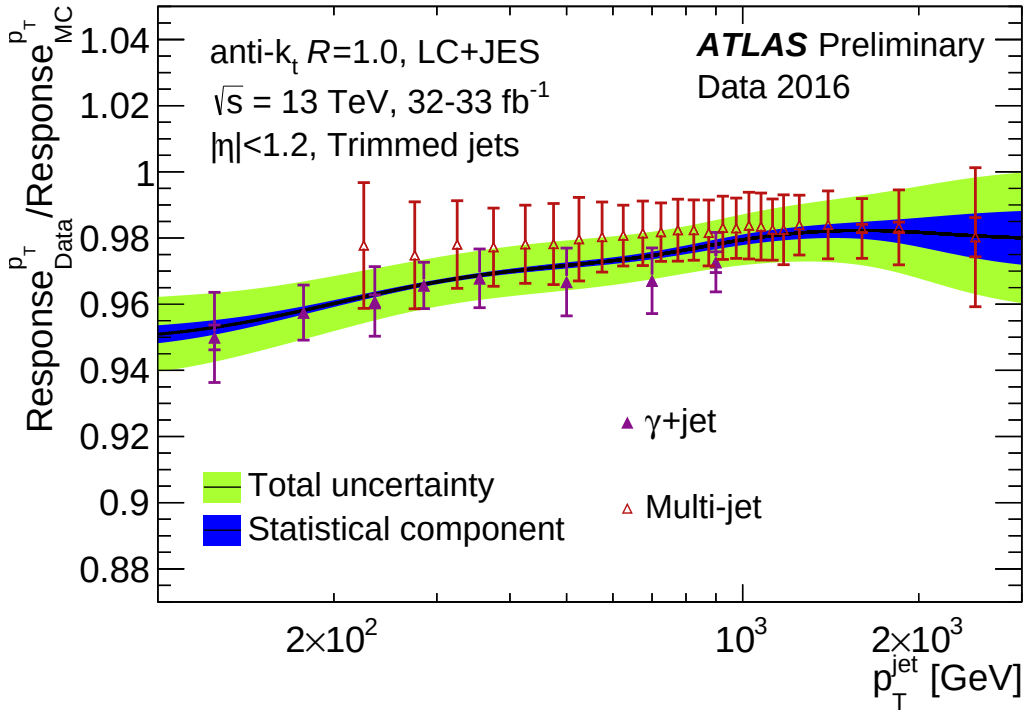


FIGURE 5.14: Average jet p_T response ratio of data-to-Monte Carlo simulation as a function of jet p_T. The combined result (blue and green bands) is based on two *in-situ* techniques: γ–jet balance method (purple markers) and the multi-jet balance (open red markers). The errors represent the statistical (inner error bars) and the total uncertainty (statistical and systematic uncertainties added in quadrature, outer error bars). The results are applied to the trimmed anti-k_t jets with $R = 1.0$ [48].

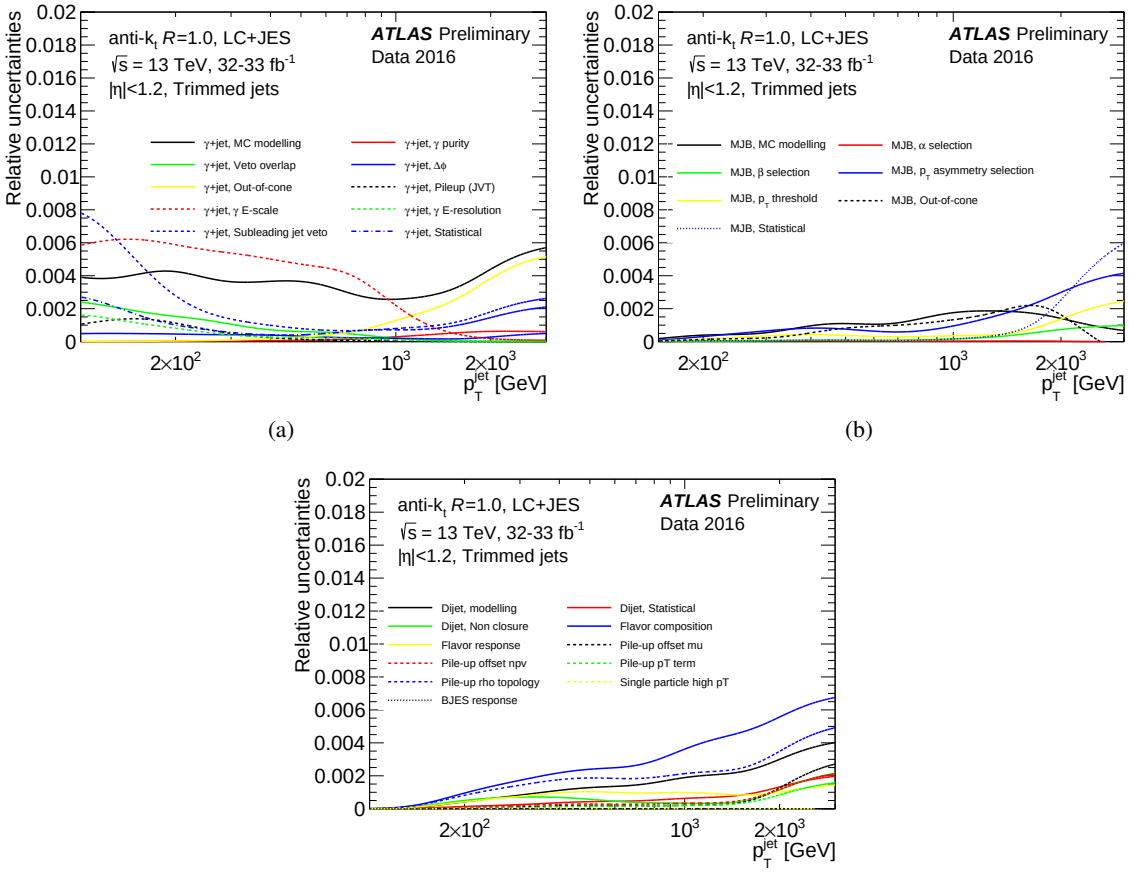
3. **Averaging:** The combination is carried out using a weighted average of the *in-situ* measurements based on a χ^2-minimisation. The weights take into account the statistical and systematic uncertainties, as well as correlations and differing bin sizes. The local χ^2 is also used to define the level of agreement between the *in-situ* measurements where they overlap.

The full set of uncertainties is propagated to the combined result using the pseudo-experiments. The sources of the systematic uncertainties are treated according to the Hessian formalism; each uncertainty source is fully correlated across the kinematic region(p_T,η), while it is assumed to be uncorrelated with the other sources. The sources of uncertainty that affect both the small-R and large-R jet *in-situ* calibration are treated as fully correlated. The reduced χ^2 is defined by $\sqrt{\chi^2/n_{dof}}$, where n_{dof} is the number of degrees of freedom (in this case, the number of combined measurements contributing to the average in a particular p_T bin). If there is disagreement between different *in-situ* measurements (i.e. when the reduced χ^2 value is larger than 1), the uncertainty sources are rescaled by the reduced χ^2. A smoothing procedure, using a variable-size sliding interval with a Gaussian function, is applied to the nominal data-to-MC response ratio and the systematic uncertainties. It removes spikes due to statistical fluctuations in the measurements, as well as discontinuities at the first and last point in a given measurement.

In Figure 5.14, the ratio of the jet p_T response in data and MC simulation is shown as a function of the jet p_T. The combined results are obtained from the two methods, namely the γ–jet and multi-jet balance methods. The two methods show some tension in the p_T range where both have results, which extends from 300 GeV up to approximately 1 TeV. In that whole range the local χ^2 per degree of freedom of the fit varies between 1.4 and 2. Below 300 GeV the only constraint comes from the γ–jet method. At higher p_T the multi-jet balance method quickly acquires more weight in the fit and extends the constraint up to well over 2 TeV.

The response in data determined by the combination of these two *in-situ* methods is 5% lower than the expectation n MC simulation at low p_T. This is significant considering the total uncertainty approaches 1%. For momenta around 1 TeV the data/MC response ratio reaches a plateau at a value about 0.98 with an uncertainty of 1%. Above 1.5 TeV the uncertainty steadily grows, and reaches 3% at 2.4 TeV.

A breakdown of the total JES uncertainty is presented graphically in Figure 5.15. Each curve corresponds to the contribution of a nuisance parameter from the small-R analysis to the large-R combination. Uncertainties on



(a) (b)

(c)

FIGURE 5.15: Breakdown of the combined uncertainty on the large-R jet p_{T} response, as a function of jet p_{T} [48]. Contributions from the nuisance parameters of the γ–jet and multi-jet balance methods are shown. As the multi-jet balance method relies on the small-R jet p_{T}, the nuisance parameters from all the methods for the small-R jet calibration (the Z–jet, the γ–jet, the multi-jet balance method) appear. The y-axis reflects not only the uncertainty introduced by a given nuisance parameter, but also the weight of the corresponding method in the combination. The nuisance parameters related to the γ–jet method (both directly and through their effect on the multi-jet balance) are shown in Figure (a). those of the multi-jet balance method are shown in Figure (b). The nuisance parameters related to the flavour, pile-up, dijet calibration, and high p_{T} particles uncertainties are shown in Figure (c).

the photon energy response are dominant in the γ–jet method at the intermediate p_{T} range, as expected. Modelling uncertainties play an important role at low p_{T}. We use a similar γ–jet method to calibrate the small-R jets, thus the γ–jet uncertainties also affect the multi-jet balance, through their contribution to the small-R JES, used as the reference scale in that method. Both contributions; (1) the large-R γ–jet uncertainties and (2) the small-R γ–jet uncertainties propagated through the multi-jet balance, are summed in Figure 5.15(a). Figures5.15(b) and 5.15(c) present the remaining nuisance parameters of the multi-jet balance method. Their contribution to the combined uncertainty is significant only at high p_{T}, where the multi-jet balance has a considerable weight in the combination. The dominant sources of the systematic uncertainties are the electron energy scale, the flavour composition and the variations of the p_{T} asymmetry cut.

Chapter 6

Event selection

In this chapter, we overview the event selections used for the search of the diboson resonances. We categorise the events depending on the production mechanism and the event topology.

- Production mechanism:
  - Merged category: Hadronically decaying vector bosons are reconstructed as large-R jets. This category shows high sensitivity in the higher mass region.
  - Resolved category: Hadronically decaying vector bosons are reconstructed pairs of small-R jets. This category is used to reconstruct the diboson resonances in the lower mass region.
- Event topology:
  - VBF category: High-p_{T} jets are observed in the forward direction. These jets are considered to be produced via the vector boson fusion process.
  - ggF/$q\bar{q}$ category: The events failing the VBF category are processed in the ggF/$q\bar{q}$ category.

A schematic summary of the event selection is shown in Figure 6.1. We describe the event selection in each of the categories in this chapter. The event selections both for the merged and resolved analyses are optimised using MC simulation samples, assuming the integrated luminosity to be 36.1 fb^{-1}.

6.1 Trigger selection

The single-lepton trigger is a straight-forward choice to collect the dataset for this analysis. We adopt a set of un-prescaled[1] triggers with the lowest threshold in each data-taking period, as summarised in Table 6.1.

6.1.1 Electron channel

To record events in the electron channel, we exploit single-electron triggers with the transverse energy thresholds of 24 GeV and 26 GeV in 2015 and 2016, respectively. These lowest threshold triggers require electron candidates

[1] Prescaled triggers randomly reject a fixed fraction of triggers to keep the total trigger rate within an allocated value

Table 6.1: The list of triggers used in the analysis. In the table, ID and ISO represent the identification and isolation criteria, which are explained in Section 4.2.4 and 4.2.5, respectively.

Data-taking period	Electron channel	Muon channel	
		p_{T} ($\mu\nu$) < 150 GeV	p_{T} ($\mu\nu$) > 150 GeV
2015	E_{T}^{e} > 24GeV, *Medium* ID E_{T}^{e} > 60GeV, *Medium* ID E_{T}^{e} > 120GeV, *Loose* ID	p_{T}^{μ} > 20GeV, *Loose* ISO p_{T}^{μ} > 50GeV	$E_{\mathrm{T}}^{\mathrm{miss}}$ > 70GeV
2016a (run < 302919) ($L < 1.0 \times 10^{34}$ cm^{-2} s^{-1})	E_{T}^{e} > 26GeV, *Tight* ID, *Loose* track ISO E_{T}^{e} > 60GeV, *Medium* ID, no d_0 cut E_{T}^{e} > 140GeV, *Loose* ID, no d_0 cut	p_{T}^{μ} > 26GeV *Medium* track ISO p_{T}^{μ} > 50GeV	$E_{\mathrm{T}}^{\mathrm{miss}}$ > 90GeV
2016b (run ≥ 302919) ($L < 1.7 \times 10^{34}$ cm^{-2} s^{-1})	same as above	same as above	$E_{\mathrm{T}}^{\mathrm{miss}}$ > 110GeV

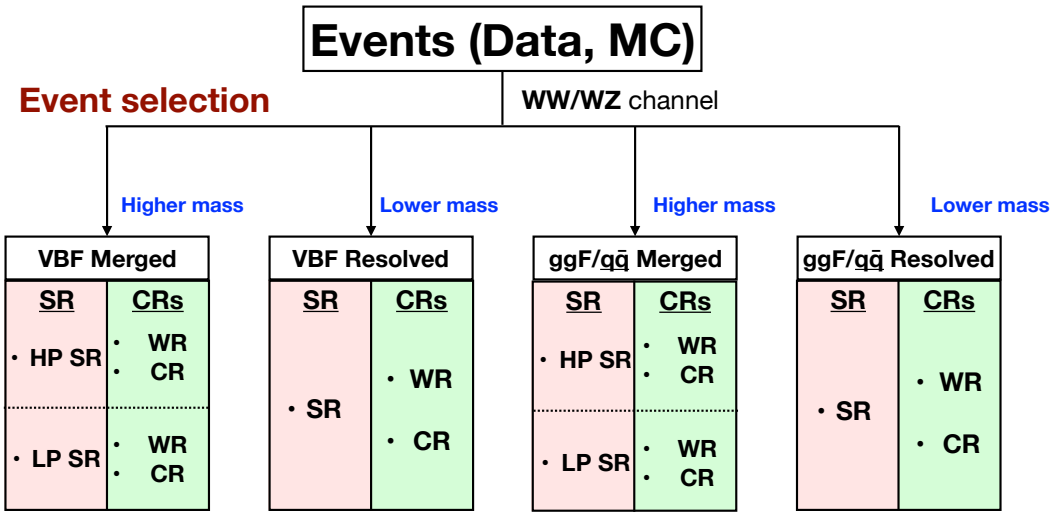


FIGURE 6.1: Schematic summary of the event categorisation and the event selections. Signal region (SR) is the main analysis region; Control region (CR) is the background-rich region, which is used to normalise the major contributions from the SM background.

to pass the isolation requirements, which have at least 90% efficiency with respect to each selected offline electron depending on the lepton p_{T}. We apply no isolation requirements, but looser identification requirements in the triggers with higher thresholds. To mitigate the efficiency loss due to the isolation requirements, a logical OR with the higher threshold triggers is used in the electron channel. Figure 6.2(a) shows the trigger efficiency taking the logical OR among the three triggers; Figure 6.2(b) shows the trigger efficiency of the lowest threshold trigger.

6.1.2 Muon channel

The single-muon triggers with the lowest p_{T} threshold of 20 (26) GeV are used to the data in 2015 (2016), whereas they are subject to a large inefficiency due to limited trigger coverage. By exploiting the single-muon triggers, about 70 % efficiency is expected in the muon channel. This efficiency is lower than the efficiency of the single-electron trigger; about 90 % is expected in the electron channel. To compensate the poor efficiency in the muon channel, we adopt the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) trigger at high-$p_{\mathrm{T}}$ region. The $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger has an online threshold of 70 GeV for the 2015 data and of 90–110 GeV for the 2016 data. The muon track p_{T} is not considered in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation in the trigger algorithm and p_{T} of the muon-neutrino system (p_{T} ($\mu\nu$)) is comparable with the $E_{\mathrm{T}}^{\mathrm{miss}}$ at the trigger level.

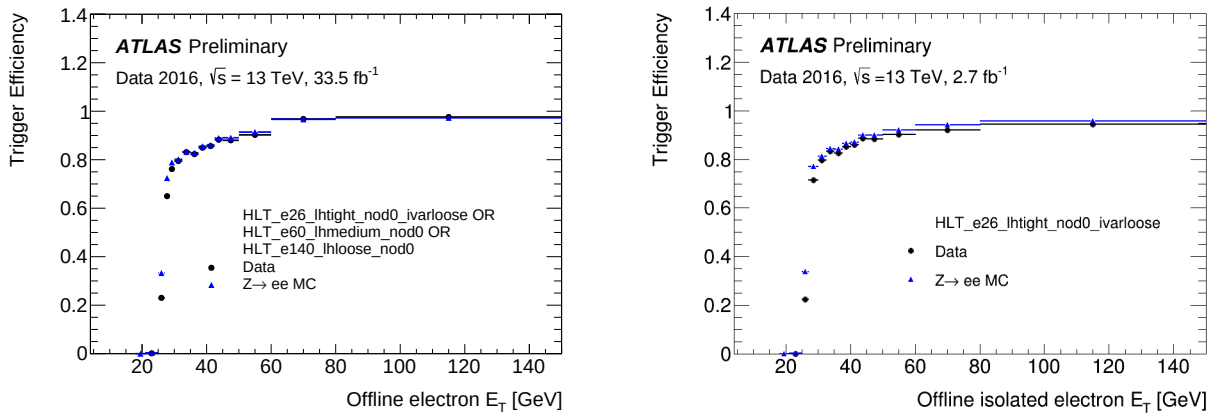
Study of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency

The performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger as a function of $p_{\mathrm{T}}(\mu\nu)$ has been studied both in the merged and resolved analyses. Since the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers are independent from the single-muon triggers, the performance is estimated using datasets collected with the single-muon triggers. For the merged category, the following selections are applied:

- exactly one single-muon with $p_{\mathrm{T}} > 27$ GeV,
- $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV,
- at least one large-R jet with $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.0$,
- passed the single-muon triggers.

For the resolved category, the following selections are applied:

- exactly one signal muon with $p_{\mathrm{T}} > 27$ GeV,
- $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV,



(a) (b)

FIGURE 6.2: (a) Trigger efficiency with taking logical OR among three single-electron triggers, as a function of transverse energy of offline electron candidates (E_T). (b) Trigger efficiency of the trigger with the lowest threshold (HLT_e26_lhtight_nod0_ivarloose) as a function of transverse energy of offline electron candidates [142]. The efficiency of the trigger with the lowest threshold in (b) is slightly lower than the efficiency of the logical OR among the three single-electron triggers in (a) due to the isolation requirements imposed on the lowest threshold trigger.

- at least two small-R jets with $p_T > 20$ GeV and $|\eta| < 2.5$,
- passed the single-muon triggers.

The E_T^{miss} trigger efficiency is evaluated as:

$$\text{Efficiency} = \frac{\text{Number of events passingthe selections and } E_T^{miss} \text{ triggers}}{\text{Number of events passing the selections}} \quad . \tag{6.1}$$

Figures 6.3(a) and 6.3(b) show the estimated efficiency of the E_T^{miss} triggers to the W+jets MC sample in the merged and the resolved category, respectively. We derived the scale factors (SF) to the events with $150 < p_T(\mu\nu) < 200$ GeV to take into account for the difference on the efficiency between the data and the W+ jets MC sample. The definition of the SF is:

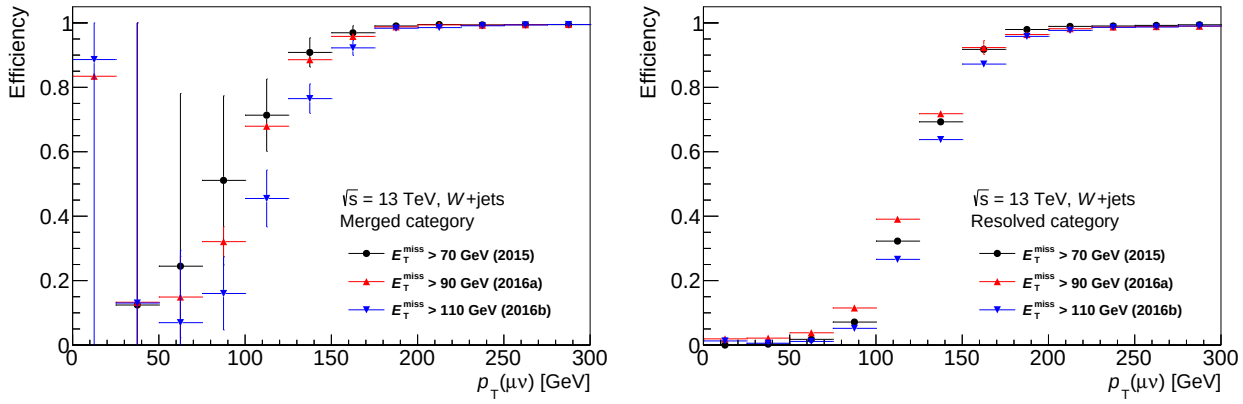
$$\text{SF} = \frac{\text{Efficiency}^{\text{Data}}}{\text{Efficiency}^{\text{MC}}} \quad . \tag{6.2}$$

The calculated SF is fitted using a error function:

$$\text{SF} = 0.5 \times \left(1 + \text{erf}\left(\frac{E_T^{miss} - p_0}{\sqrt{2} \times p_1}\right)\right) \quad , \tag{6.3}$$

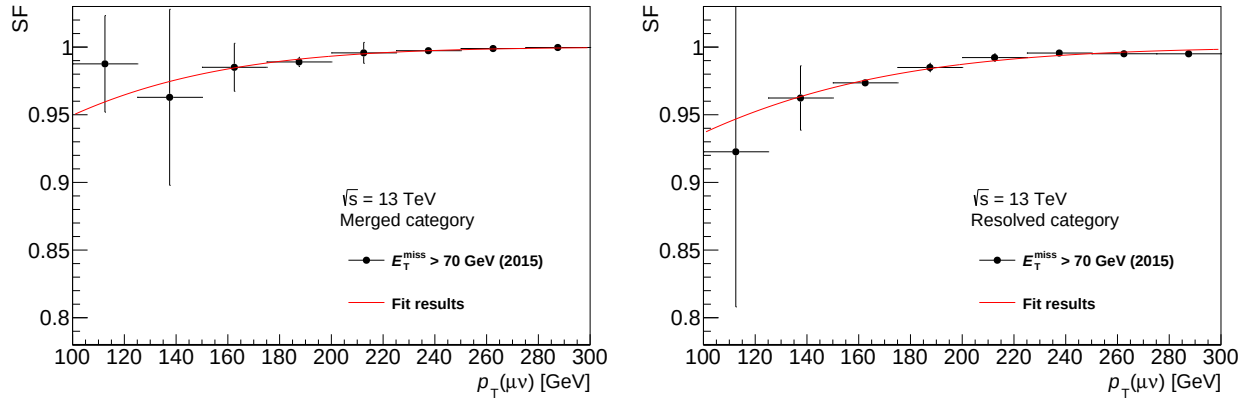
where p_0 and p_1 represent the threshold and the width of the fit function. The systematic uncertainty on the scale factors are evaluated to be approximately 2% by comparing the results with the results obtained by the $t\bar{t}$ MC sample.

We expect approximately 100 % efficiency at the region of $p_T(\mu\nu) > 200$ GeV. In the merged analysis, we only use the E_T^{miss} triggers in the muon channel since we apply an offline selection with $p_T(\ell\nu) > 200$ GeV. The signal acceptance improves by about 30 % using the E_T^{miss} triggers instead of the single-muon triggers; by this improvement, the sensitivity in the combined $e + \mu$ channel improves by about 10 % in particular at the high-mass region, where the sensitivity is mainly dependent on the signal acceptance. In the resolved analysis, the E_T^{miss} triggers are used at p_T $(\mu\nu) > 150$ GeV and the events with p_T $(\mu\nu) < 150$ GeV rely on the single-muon triggers. With the steady increment in the LHC luminosity, the trigger menu is kept updated to suppress the trigger rate below the budget. To follow the changes on the trigger menu, thresholds on the E_T^{miss} triggers are changed for different periods of data-taking as shown in Table 6.1.



(a) (b)

FIGURE 6.3: (a) efficiency of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger to the W+jets MC sample as a function of $p_{\mathrm{T}}(\mu\nu)$ in the merged category, and (b) in the resolved category. Three different thresholds were applied to the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers according to data-taking periods as shown in Table 6.1.



(a) (b)

FIGURE 6.4: (a) Scale factors to the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger for the 2015 dataset as a function of $p_{\mathrm{T}}(\mu\nu)$ in the merged category, and (b) in the resolved category. The error bars represent the statistical uncertainties.

TABLE 6.2: Selections for (a) the electron, (b) muon, (c) small-R jet, and (d) large-R jet candidates used in this analysis. Veto and signal leptons are defined to require exactly one lepton in a event.

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

(a)

	Muons	
	Veto	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	< 3
$\|z_0 \sin\theta\|$	< 0.5 mm	< 0.5 mm

(b)

	Signal Small-R Jets
Algorithm	anti$-k_t$, $R = 0.4$
p_T threshold	20 GeV
$\|\eta\|$	< 2.5 (4.5 for VBF jets)
Quality	not "bad" jet
Pile-up Removal	JVT < 0.59 when $\|\eta\| < 2.4$ && $p_T < 60$ GeV
b-Tagging (if applied)	`MV2c10`, 85% efficiency

(c)

	Signal Large-R Jets
Algorithm	anti$-k_t$, $R = 1.0$
p_{T} threshold	200 GeV
$\|\eta\|$	< 2.0
Mass threshold	50 GeV

(d)

6.2 Pre-selections of event

The selections for each of the object candidates used in this analysis are summarised in the Table 6.2. We applied the following selections as 'pre-selection' to the recorded events:

Good Runs List (GRL)

To assure a good data quality, the following requirements are applied as the Good Runs List:

- LHC was in the stable-beam mode,
- Magnets were ON,
- All sub-detectors were working without any errors.

Primary vertex

The events are required to have at least one primary vertex: a vertex with at least two tracks with $p_{\mathrm{T,trk}} > 400\,\mathrm{MeV}$, where $p_{\mathrm{T,trk}}$ is the transverse momentum of tracks associated with the vertex.

Sub-detector error veto
Events are rejected if any of the sub-detectors was in an error state.

Other selections
To reduce the data size, the candidate events are required to satisfy the following requirements:

- One signal lepton is reconstructed,
- There is no additional "veto" leptons,
- $p_{\mathrm{T}}(\ell\nu) > 75\,\mathrm{GeV}$,
- At least one large-R jet or at least two small-R jets are reconstructed

6.3 Event categorisation

The sensitivity to resonances of different masses is optimised by categorising the events according to the topology, production mechanism, and amount of background. This section provides the description for each of the categories. The event selection criteria are summarised in Table 6.3 and 6.4 for the merged and resolved analyses respectively. Figure 6.5 illustrates the jet selections used to reconstruct the hadronically decaying V boson candidates in the signal and control regions of the analysis. The mass of either the large-R jet (m_{J}) or the system of two small-R jets (m_{jj}) is used to define the "mass windows", as shown in Figure 6.5.

6.3.1 Merged category

Taking into account the different bosons in the signal assumption, two sub-channels are considered. These channels are defined based on the tagging of the large-R jet and are namely:

- WZ channel: The large-R jet is tagged as a Z boson,
- WW channel: The large-R jet is tagged as a W boson.

There is a large overlap between these sub-channels because the mass of W and Z bosons are close to each other and the resolution of the jet mass measurement in the tagger is wide compared to the mass difference. In the following section, the optimisation of the selection cuts is performed in the combined WW+WZ channel, where the large-R jet is required to be tagged as either W or Z boson.

High-purity signal region

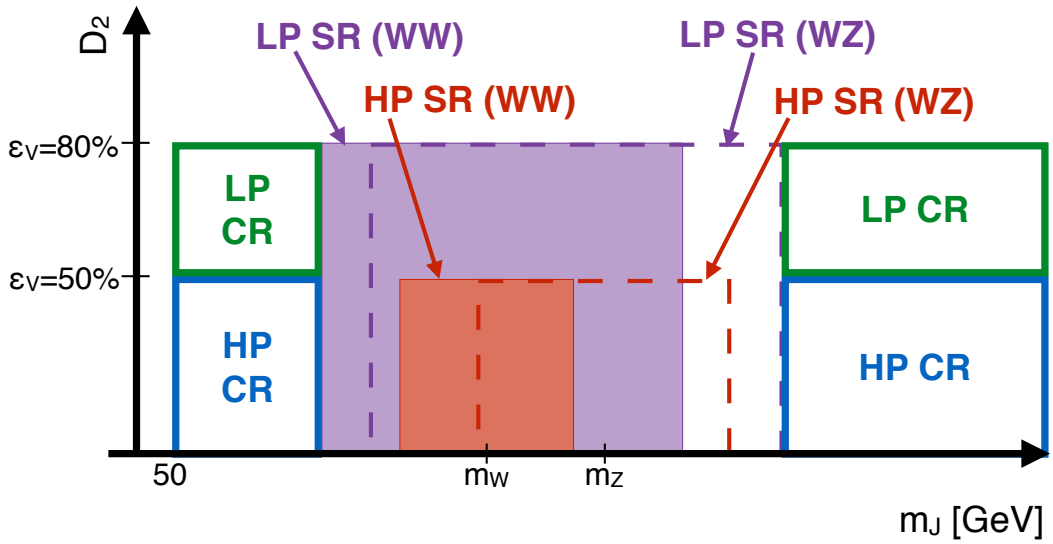
High-purity (HP) signal region (SR) is the main channel of the analysis showing the maximum sensitivity to the signals at high-mass region. The event selections for the HP signal region are optimised using the HVT signals.

Selections on $E_{\mathrm{T}}^{\mathrm{miss}}$
$E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be greater than 100 GeV. This cut is applied to suppress the multi-jet background. After this selection, there still remains the multi-jet background especially in the electron channel. In order to suppress the remaining multi-jet background, we introduced a further event cleaning using $E_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}(\ell\nu)$. Figures 6.6(a) and 6.6(b) show the $E_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}(\ell\nu)$ distributions in the electron and the muon channels, respectively. From these distributions,

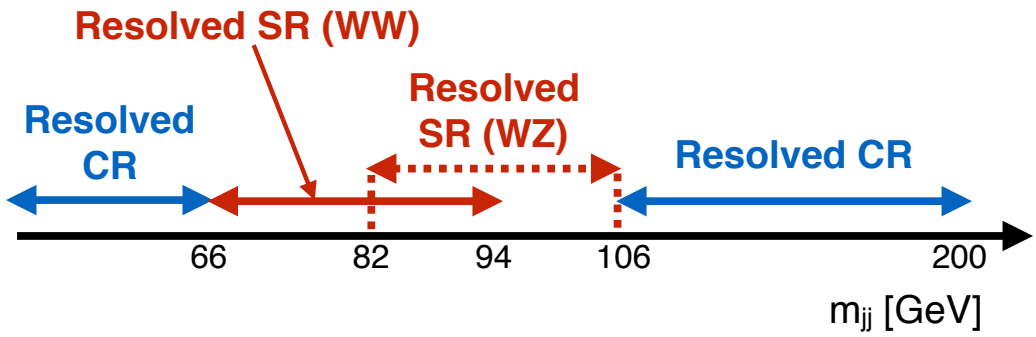
$$\frac{E_{\mathrm{T}}^{\mathrm{miss}}}{p_{\mathrm{T}}\,(e\nu)} > 0.2 \tag{6.4}$$

is required only in the electron channel. The $E_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}(\ell\nu) < 0.2$ selection, however, tends to reject right-handed W^+ bosons. Figure 6.7(a) shows $\cos\theta_{2\mathrm{D}}$ distributions in the W+jets sample. The "transverse helicity" angle $\theta_{2\mathrm{D}}$ is defined as:

$$\cos\theta_{2\mathrm{D}} = \frac{\overrightarrow{p_{\mathrm{T}}^{\ell*}} \cdot \overrightarrow{p_{\mathrm{T}}^{W}}}{|\overrightarrow{p_{\mathrm{T}}^{\ell*}}||\overrightarrow{p_{\mathrm{T}}^{W}}|} \quad , \tag{6.5}$$

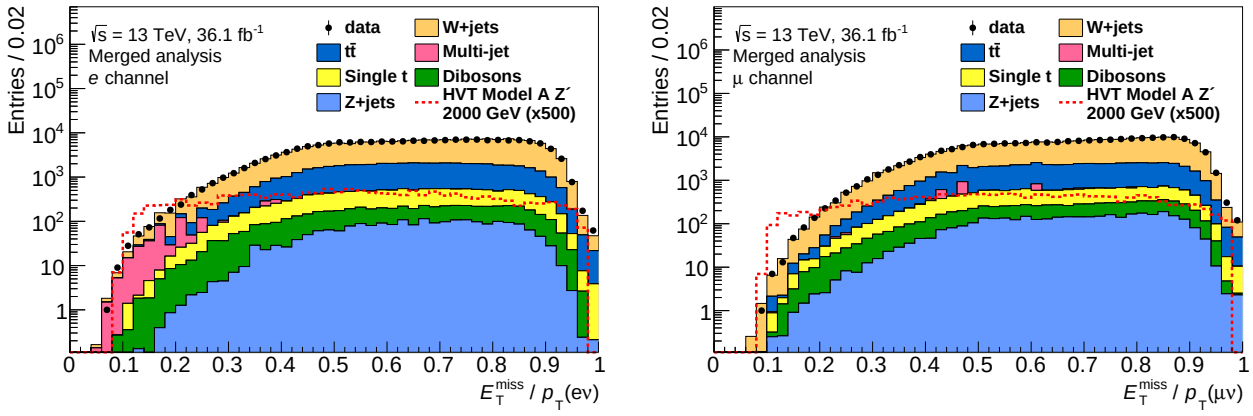


(a)



(b)

FIGURE 6.5: (a) Illustration of the merged WW (shaded area) and WZ (dashed lines) signal regions (SR) according to the large-R jets selection. The 50% and 80% V-tagging efficiency (ε_V) working points are used to form the high-purity (HP) and low-purity (LP) regions respectively. For each working point, a jet mass requirement is imposed and an upper bound on the substructure variable D_2 is set. Since both requirements depend on the p_T of the large-R jet, an absolute definition is not given in the figure. (b) Definitions of the resolved WW and WZ SR based on the dijet mass selection. In both channels, the SR mass sidebands are used to define the W+jets control region (CR) [38].



(a) (b)

FIGURE 6.6: (a) $E_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}(\ell\nu)$ distribution in the electron channel and (b) in the muon channel. The observed data, a set of expected background samples, and a signal prediction are super imposed. To suppress the multi-jet background in the electron channel, we introduced a selection of $E_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}(\ell\nu) < 0.2$.

where $\overrightarrow{p_{\mathrm{T}}^{\ell*}}$ is the transverse momentum of the lepton in the transverse W rest frame and $\overrightarrow{p_{\mathrm{T}}^{W}}$ is the transverse momentum of the W boson in the laboratory frame [143]. Each template distribution of left-handed, longitudinal, and right-handed W^+ bosons are shown in Figure 6.7(b). From Figure 6.7(a), it is confirmed that less than 7% of right-handed W^+ bosons are rejected by this selection. Since we do not apply the $E_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}(\ell\nu) < 0.2$ selection in the muon channel, we will observe flavour asymmetries to the signals with transversely polarised W bosons.

Selections on large-R jets

In addition to the pre-selections, at least one signal large-R jet is required according to the event topology. The jet with the highest p_{T} (leading jet) is used in the analysis flow if there are multi large-R jets in a event. We use the working point of the boson tagger with 50 % efficiency in the HP SR. Furthermore, a relative boson p_{T} cut is introduced to enhance signals:

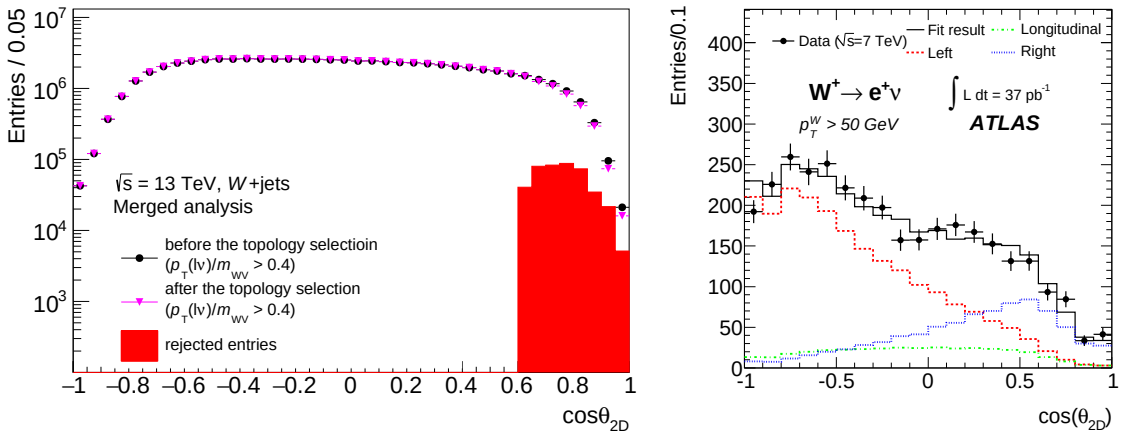
$$R = \frac{p_{\mathrm{T}}(\mathrm{V})}{m_{\mathrm{WV}}} > 0.4 \quad , \tag{6.6}$$

where $p_{\mathrm{T}}(V)$ ($V = J$ or $\ell\nu$) is the transverse momentum of the vector bosons reconstructed as the large-R jet or the lepton-neutrino system. Figure 6.8 shows the $p_{\mathrm{T}}(\mathrm{V})/m_{\mathrm{WV}}$ distributions of the observed data, expected background samples, and a signal sample. The optimal threshold on $p_{\mathrm{T}}(\mathrm{V})/m_{\mathrm{WV}}$ does not depend on signal mass point and signal models (spin-0, -1, and -2).

b-jet veto

After the kinematic and event topology cuts, main contributions from the SM background remaining in the signal region are top quark pair production ($t\bar{t}$, 54%) and W boson production associated with hadron jets (W+jets, 34%). For further rejection of the $t\bar{t}$ background, a b-jet veto is applied in the event. If there is at least one b-jet with $\Delta R > 1.0$ from the large-R jet axis, the event is removed. This cut rejects ~70 % of $t\bar{t}$ background in the SR while signal efficiency is approximately 95 %, not depending on the signal mass.

Finally, the background components in the SR are $t\bar{t}$ and W+jets with the contribution of 32 % and 55 %, respectively. These two main sources of the background are constrained in a pair of dedicated control regions (CR). The CRs are defined as kinematically close to the SR as possible, in order to reduce the uncertainty on the extrapolation. The selection cuts used to define the SR and CR are summarised in Table 6.3.



(a) (b)

FIGURE 6.7: (a) $\cos\theta_{2D}$ distribution in the W+jets sample before and after the topology selection ($E_T^{miss}/p_T(\ell\nu) < 0.2$) in the merged analysis. Events are rejected in the high $\cos\theta_{2D}$ region ($0.65 < \cos\theta_{2D}$), where most of W^+ bosons are right-handed. (b) $\cos\theta_{2D}$ distributions of each template of left-handed (red), longitudinal (green), and right-handed (blue) W^+ bosons [143].

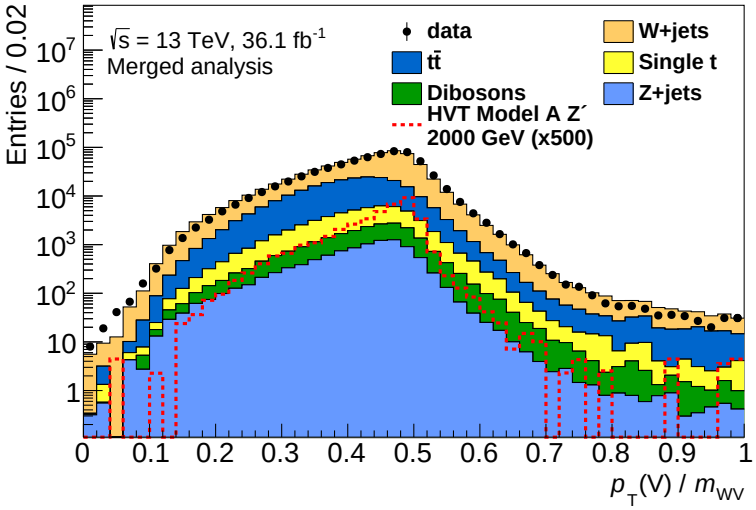


FIGURE 6.8: $p_T(V)/m_{WV}$ distributions of the observed data, expected background samples, and a HVT Z' signal sample. $p_T(V)/m_{WV} > 0.4$ is applied in this analysis.

W+jets Control Region

The W+jets control region (WR) is defined using the boson taggers with 50 % and 80 % efficiency; this definition is schematically shown in Figure 6.5(a). In the WR, we require the large-R jet to fail the selection on the jet mass providing 80 % efficiency for W and Z, while passing the selection on the substructure variable D_2 providing 50 % efficiency for W and Z. For the simplicity of the analysis, we use a common control region in all of the WW, WZ, and combined channels. The purity of the W+jets background is 77 % in the lower mass side-band, 62 % in the higher mass side-band region and 65 % in total. The contamination from the $t\bar{t}$ background is 30 % in the higher mass side band region.

$t\bar{t}$ Control Region

The $t\bar{t}$ control region (TR) is defined by requiring at least one b-jet instead of vetoing b-jets. The purity of the $t\bar{t}$ background in this region is 85 %, while the signal leakage in both W+jets and $t\bar{t}$ control regions is negligible. The purity of W+jets in the WR is slightly poor, but the simultaneous fit to the TR and WR makes it possible to determine the normalisation of W+jets correctly, thanks to high purity of $t\bar{t}$ in the TR.

Low-Purity selection

Since we use the working point of the boson tagger with 50 % efficiency in the HP SR, many signal events leak to the high-D_2 region where a low signal-to-background ratio is expected. To improve the signal acceptance, in particular in the high-mass region, the Low-Purity (LP) SR is introduced. Events failing either the D_2 cut or the mass window cut at the working point of the boson tagger with 50 % efficiency, can be recovered in this SR if the events are accepted by the working point of the boson tagger with 80 % efficiency. The other selection cuts are the exactly same as those for the HP SR, as summarised in Table 6.3.

The background components are 72 % from W+jets and 19 % from $t\bar{t}$ in the LP SR. Dedicated CRs for the LP SR (both WR and TR) are defined with the 80 % working point cuts as well. The LP WR requires events to fail the 80 % mass cut, to fail the 50 % D_2 cut, and to pass the 80 % D_2 cut. The LP TR requires events to fail either the 50 % D_2 or mass cut, to pass both the 80 % D_2 and mass cut, and to have at least one b-jet. The purity of W+jets ($t\bar{t}$) in WR (TR) is 68 % (80 %). Since the large-R jet in the $t\bar{t}$ background mainly comes from the hadronically-decaying W boson, the fractions of the $t\bar{t}$ background in the SR and WR are suppressed by the inverted D_2 cut. The simultaneous fit to HP and LP SRs improves the sensitivity by 10 %.

6.3.2 Resolved category

We can select $W \to \ell\nu$ candidates in the same criteria as the merged analysis but requiring $p_{\mathrm{T}}(\ell\nu) > 75$GeV. The requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ is loosened to $E_{\mathrm{T}}^{\mathrm{miss}} > 60\,\mathrm{GeV}$ for the acceptance to the low-mass signals ($m < 500\,\mathrm{GeV}$). Therefore, non-negligible contribution from the multi-jet background was expected in the resolved analysis. We performed a data-driven estimation of the multi-jet background. (The details are described in Section 7.1.) However, expected contribution of the multi-jet background is less than 10 % in the signal regions, thus we neglected the effect of the multi-jet contribution for the optimisation of the selection cuts.

We select a $W \to jj$ candidate from a pair of the small-R jets which have the highest and second-highest p_{T} in the event, applying the following requirements; $p_{\mathrm{T}} > 45\,\mathrm{GeV}$ and $|\eta| < 2.4$. At least one of the selected candidates is required to satisfy $p_{\mathrm{T}} > 60\,\mathrm{GeV}$. To take into account the large branching fraction of $W \to c\bar{s}$ and $Z \to b\bar{b}$, we allow up to one b-jet in a $W \to jj$ candidate or up to two b-jets in a $Z \to jj$ candidate. If we observe at least one b-jet not associated to W or Z, the event is rejected. The requirement for the m_{jj} to be consistent with either the W or Z mass is optimised using the HVT W' and HVT Z' samples, letting the high-mass and low-mass thresholds be optimised simultaneously. Her, the optimised parameter is asymptotic sensitivity:

$$\sigma^2 = \sum_i^{N_{\mathrm{bin}}} \left(\frac{s_i}{s_i + b_i + (\Delta b_i)^2} \right)^2 \quad , \tag{6.7}$$

where s_i and b_i are expected number of signals and background in each i bin of the m_{WV} distribution. We use $N_{\mathrm{bin}} = 30$ in the range of $0 < m_{WV} < 1.5\,\mathrm{TeV}$. Δ indicates an assumed systematic uncertainty on the background

TABLE 6.3: Summary of the selection criteria used to define the merged WW and WZ signal regions (SR) and their corresponding W+jets control regions (WR) and $t\bar{t}$ control regions (TR) in the high-purity (HP) and low-purity (LP) categories. The events are also categorised with their production mechanism; the VBF selection is prioritised and the remaining events are assigned to the ggF/$q\bar{q}$ category [38].

Selection		SR: HP (LP)	WR: HP (LP)	TR: HP (LP)
Production category	VBF	$m^{\text{tag}}(j,j) > 770\,\text{GeV}$ and $\lvert\Delta\eta^{\text{tag}}(j,j)\rvert > 4.7$		
	ggF/$q\bar{q}$	Fails VBF selection		
$W \to \ell\nu$ selection	Num. of signal leptons	1		
	Num. of veto leptons	0		
	$E_{\text{T}}^{\text{miss}}$	$> 100\,\text{GeV}$		
	$p_{\text{T}}(\ell\nu)$	$> 200\,\text{GeV}$		
	$E_{\text{T}}^{\text{miss}}/p_{\text{T}}(e\nu)$	> 0.2		
$V \to J$ selection	Num. of large-R jets	≥ 1		
	D_2 eff. working point (%)	Pass 50 (80)	Pass 50 (80)	Pass 50 (80)
	Mass window Eff. working point (%)	Pass 50 (80)	Fail 80 (80)	Pass 50 (80)
Topology criteria	$p_{\text{T}}(\ell\nu)/m(WV)$ $p_{\text{T}}(J)/m(WV)$	> 0.3 for VBF and > 0.4 for ggF/$q\bar{q}$ category		
Num. of b-tagged jet	excluding b-tagged jets with $\Delta R(J,b) \leq 1.0$	0		≥ 1

estimation, and we assign $\Delta = 1.5\,\%$ here. For mass ranges of interest for this channel, the optimised W window is $66 < m_{jj} < 94\,\text{GeV}$ and the Z mass window is $82 < m_{jj} < 106\,\text{GeV}$.

The following selection cuts are applied to take the event topology of which two high-p_{T} bosons are back–to–back in the x–y plane:

- $\Delta\phi(\ell, E_{\text{T}}^{\text{miss}}) < 1.5$,
- $\Delta\phi(j1, j2) < 1.5$,
- $\Delta\phi(\ell, j1(2)) > 1.0$,
- $\Delta\phi(j1(2), E_{\text{T}}^{\text{miss}}) > 1.0$,
- $p_{\text{T}}(\ell\nu)/m_{WV} > 0.3(0.35)$ for the VBF (ggF) category,
- $p_{\text{T}}(jj)/m_{WV} > 0.3(0.35)$ for the VBF (ggF) category,

where $\Delta\phi(a,b)$ is the distance in the azimuthal direction between objects a and b. A cut on the relative boson p_{T}, $R = p_{\text{T}}(V)/m_{WV}$, is also introduced to the resolved analysis. The threshold on R is optimised to maximise the asymptotic sensitivity as above. A lower R cut is found to be optimal for the $m = 300\,\text{GeV}$ signal. The optimisation is based on $m = 300\,\text{GeV}$ signal since the sensitivity at $m > 500\,\text{GeV}$ is dominated by the merged analysis. A slightly lower threshold is preferred in the VBF category ($R > 0.3$) to the ggF/$q\bar{q}$ category ($R > 0.35$). The multi-jet contribution in the optimisation is not considered because of the same reason as the mass window optimisation. The selections in the resolved analysis are summarised in Table 6.4.

Reconstruction of m_{WV}

The invariant mass of the WV system, m_{WV}, is reconstructed from a lepton, a neutrino, and a hadronically-decaying boson candidate (large-R jet or two small-R jets). The reconstruction of the momentum of the neutrino in the z-direction, p_z, is obtained by imposing a constraint on the mass of the lepton-neutrino system to be consistent with the W boson mass.

TABLE 6.4: Summary of the selection criteria in the resolved analysis for the WW and WZ signal regions (SR), W+jets control region (WR) and $t\bar{t}$ control region (TR). The events are also categorised with their production mechanism. The VBF selection is prioritised and the remaining events are assigned to the ggF/$q\bar{q}$ category [38].

Selection		WW (WZ) SR	WR	TR
Production category	VBF	$m^{\text{tag}}(j,j) > 770\,\text{GeV}$ and $\lvert\Delta\eta^{\text{tag}}(j,j)\rvert > 4.7$		
	ggF/$q\bar{q}$	Fails VBF selection		
$W \to \ell\nu$ selection	Num. of signal leptons	1		
	Num. of veto leptons	0		
	$E_{\text{T}}^{\text{miss}}$	$> 60\,\text{GeV}$		
	$p_{\text{T}}(\ell\nu)$	$> 75\,\text{GeV}$		
	$E_{\text{T}}^{\text{miss}}/p_{\text{T}}(e\nu)$	> 0.2		
$V \to jj$ selection	Num. of small-R jets	≥ 2		
	$p_{\text{T}}(j_1)$	$> 60\,\text{GeV}$		
	$p_{\text{T}}(j_2)$	$> 45\,\text{GeV}$		
	$m(jj)$ [GeV]	[66, 94] ([82, 106])	< 66 or [106, 200]	[66, 106]
Topology criteria	$\Delta\phi(j,\ell)$	> 1.0		
	$\Delta\phi(j, E_{\text{T}}^{\text{miss}})$	> 1.0		
	$\Delta\phi(j,j)$	< 1.5		
	$\Delta\phi(\ell, E_{\text{T}}^{\text{miss}})$	< 1.5		
	$p_{\text{T}}(\ell\nu)/m(WV)$ $p_{\text{T}}(jj)/m(WV)$	> 0.3 for VBF and 0.35 for ggF/$q\bar{q}$ category		
Num. of b-tagged jets	$j_1 \equiv b$ or $j_2 \equiv b$ where $V \to j_1 j_2$	$\leq 1(2)$	≤ 1	> 0 (for jets other than j_1 or j_2)
	$j_1 \neq b$ and $j_2 \neq b$ where $V \to j_1 j_2$	0		

In the resolved analysis, the m_{WV} is reconstructed by giving the W-mass constraint also on the di-jet system:

$$p_{\text{T}}^{\text{corr}} = p_{\text{T}}(jj) \times \frac{m_W}{m_{jj}} \quad , \tag{6.8}$$

$$m^{\text{corr}} = m_W \quad , \tag{6.9}$$

where m_{jj} and m_W are the invariant mass of the reconstructed hadronically-decaying W boson and the PDG value of the W-boson mass, respectively. For the WZ signal interpretation, we apply a same Z-mass constraint on the m_{jj} instead.

6.3.3 VBF category: selections of the VBF jets

We expect high-p_{T} jets in the forward direction in the signal events produced via the vector boson fusion process (VBF). Two small-R jets with the maximum m_{jj}^{VBF} in the event are selected as the candidates of the "VBF jets", where m_{jj}^{VBF} is the invariant mass reconstructed by the two jets. The VBF jets need to satisfy the following requirements; $p_{\text{T}} > 30\,\text{GeV}$, $|\eta| < 4.5$, $\eta(j1)^{\text{VBF}}\eta(j2)^{\text{VBF}} < 0$, and neither of the jets is tagged as a b-jet. To gain sensitivity in the VBF signals,

- $m_{jj}^{\text{VBF}} > 770\,\text{GeV}$ and
- $\left|\Delta\eta_{jj}^{\text{VBF}}\right| > 4.7$

are required, where $\Delta\eta_{jj}^{\text{VBF}}$ are the difference in pseudo-rapidity between the two jets. The criteria to select the VBF jets are summarised in Table 6.5.

TABLE 6.5: Summary of criteria to select VBF-jet candidates.

$N(j)$ with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $\lvert\eta\rvert < 4.5$ and not b-tagged	≥ 2 (merged), ≥ 4 (resolved)
$\eta(j1) \times \eta(j2)$	< 0
m_{jj}	$> 770\,\mathrm{GeV}$
$\Delta\eta_{jj}$	> 4.7

In the ggF/$q\bar{q}$ category, all small-R jets in $|\eta| < 2.5$ and large-R jets in $|\eta| < 2.0$ are used to select $W/Z \to qq$ candidates. On the other hand, in the VBF category, both candidates of the VBF jets should not overlap with $W/Z \to qq$ candidate. In the resolved analysis, small-R jets that satisfy the VBF selection are removed from consideration of the possible $W/Z \to qq$ candidates. Thus, the number of small-R jets in the region of $|\eta| < 4.5$ is required to be greater than 3 in the resolved analysis. In the merged analysis, $\Delta R(j, J) > 1.5$ is required when we select the large-R jet, where $\Delta R(j, J)$ is the distance between small-R and large-R jets in the η–ϕ plane. Selection efficiency of the VBF jets to the signals produced via VBF is about 28 %. About 1 % (3 %) of the ggF/$q\bar{q}$ signals migrate into the VBF category in the merged (resolved) analysis.

6.3.4 Optimisation of signal regions

Figure 6.9 shows the signal acceptance times efficiency $\epsilon \times A$ in the ggF/$q\bar{q}$ category[2]. $\epsilon \times A$ is plotted as a function of signal mass for the HVT WW, HVT WZ, RS G^*, and heavy neutral higgs signal models in all the merged and resolved signal regions, as well as the total combined acceptance. The resolved region shows stable $\epsilon \times A$ values between $8 - 12$ % up to around 800 GeV; the merged HP (LP) region has stable $\epsilon \times A$ values between $18 - 20$ % ($8 - 10$ %) above 800 GeV. The total acceptance in the merged plus resolved channels is flat above 800 GeV, ranging from $28 - 32$ %, and decreases to 16 % in the lowest mass region. The $\epsilon \times A$ values are generally lower ($6 - 12$ %) for the heavy scalar signal because the two bosons are produced less centrally than for the spin-1 and spin-2 signals, thus the $p_{\mathrm{T}}(V)/m(WV)$ requirements reject more signal.

Figure 6.10 shows the signal acceptance times efficiency in the VBF category. $\epsilon \times A$ is plotted as a function of signal mass for the HVT WW, HVT WZ, and heavy scalar signal models in all the merged and resolved signal regions, as well as the total combined acceptance. The resolved region shows the $\epsilon \times A$ values about 2% up to around 700GeV; the merged HP (LP) region has the $\epsilon \times A$ values between $3 - 6$ % ($2 - 4$ %) above 700 GeV. The total merged plus resolved acceptance is ranging from $4 - 8$ %, and decreases to 2 % in the lowest mass region. Experimental factors, such as the detector coverage and the pile-up effect, lead to low tagging efficiency of the VBF jets resulting in small $\epsilon \times A$. To increase the sensitivity, we give a priority to the VBF category because of the small signal $\epsilon \times A$ in the VBF category and the high fraction of the VBF signal that leaks in the ggF/$q\bar{q}$ category. The leakage occurs due to inefficiencies related to the reconstruction and identification of the VBF-jets, and results in a small degradation in sensitivity.

6.4 Definition of the merged and resolved analyses

To be summarised, we define 6 SRs (merged ggF HP, merged ggF LP, merged VBF HP, merged VBF LP, resolved ggF, and resolved VBF) and 12 CRs corresponding to them (TR and WR for each). Figure 6.11 summarises how to define the SR/CRs in this analysis, we give a priority to the merged SRs. The events failing the merged selection are processed in the resolved SRs. The events failing all of the SR selections are processed in the control regions.

[2] $\epsilon \times A$ is defined as the ratio of the number of signal events in the SR to the number of generated signal events.

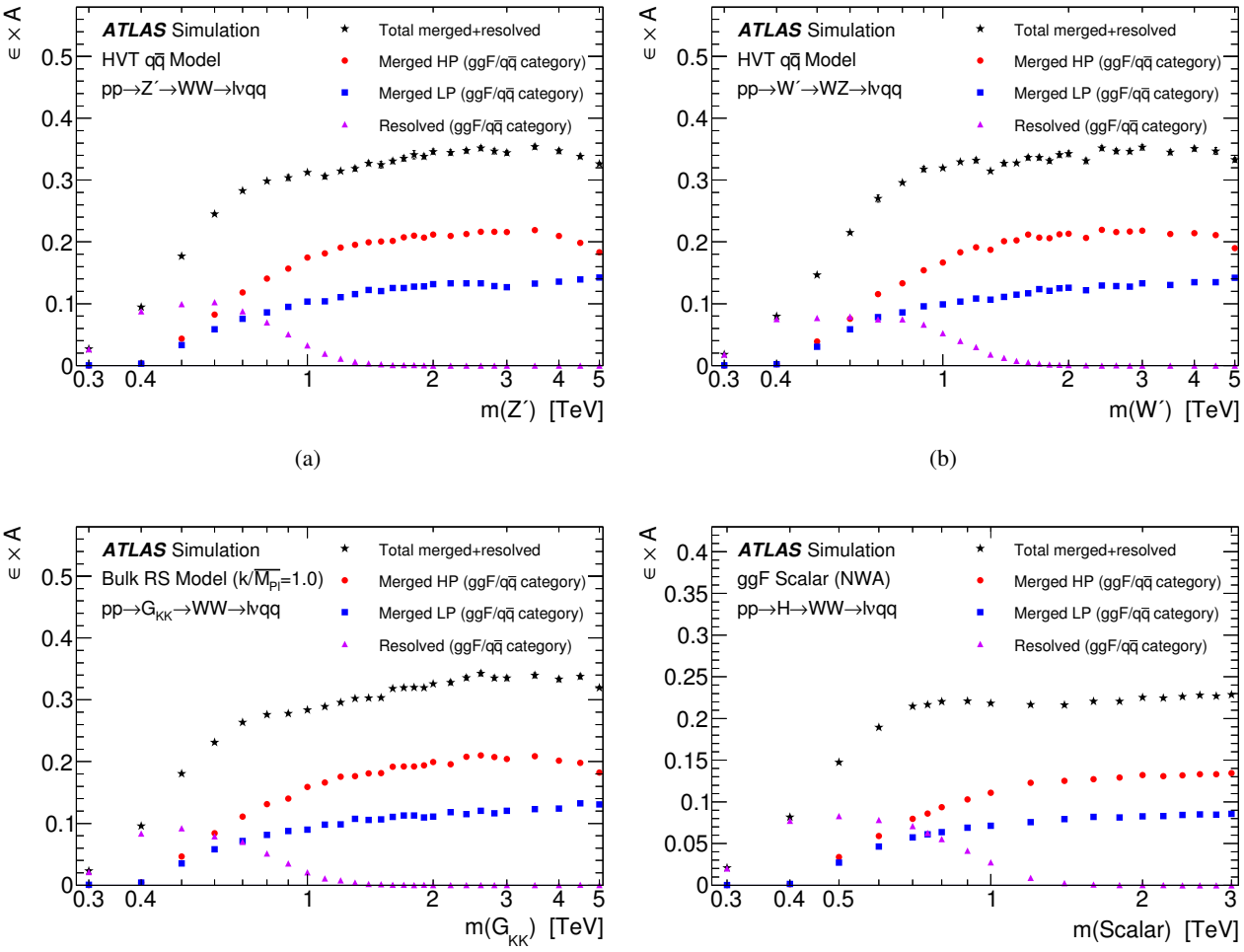


FIGURE 6.9: Product of signal efficiency (ϵ) and acceptance (A), presented in the ggF/qq̄ category for(a) HVT $WW \to \ell\nu qq$, (b) HVT $WZ \to \ell\nu qq$, (c) RS $G_{\mathrm{KK}} \to \ell\nu qq$, and (d) neutral scalar signal (H) decaying into $\ell\nu qq$ in the narrow-width approximation (NWA) in the various analysis categories. $\epsilon \times A$ is defined as the ratio of the number of signal events reconstructed in the signal region to the number of generated signal events [38].

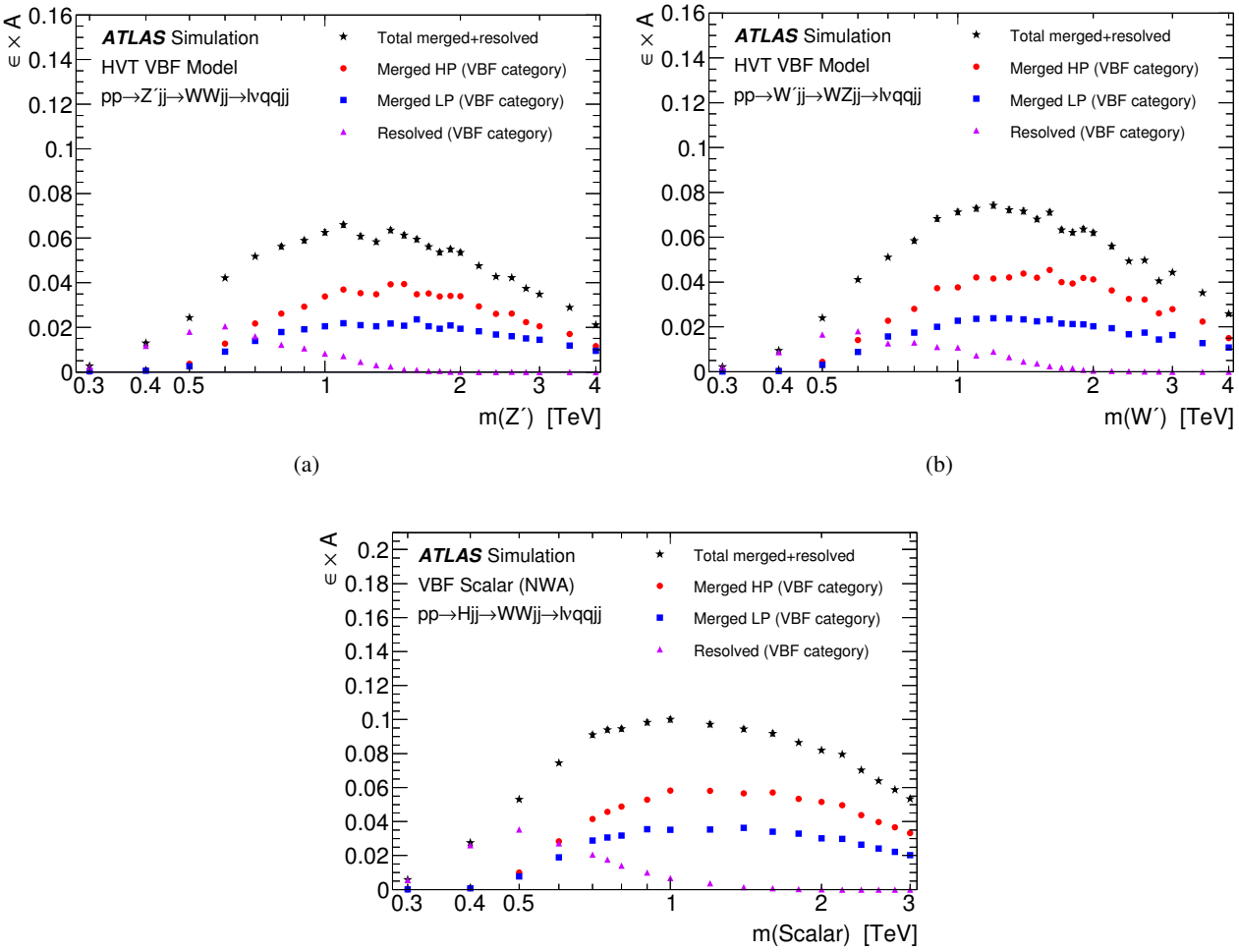


FIGURE 6.10: Product of signal efficiency (ϵ) and acceptance (A) for signals produced via the VBF mechanism, presented in both the ggF/$q\bar{q}$ and VBF categories, for (a) HVT $WW \to \ell\nu qq$, (b) HVT $WZ \to \ell\nu qq$, and (c) neutral scalar signal (H) in the narrow-width approximation (NWA) decaying into $\ell\nu qq$ in the various analysis signal regions. $\epsilon \times A$ is defined as the ratio of the number of signal events reconstructed in the signal region to the number of generated signal events [38].

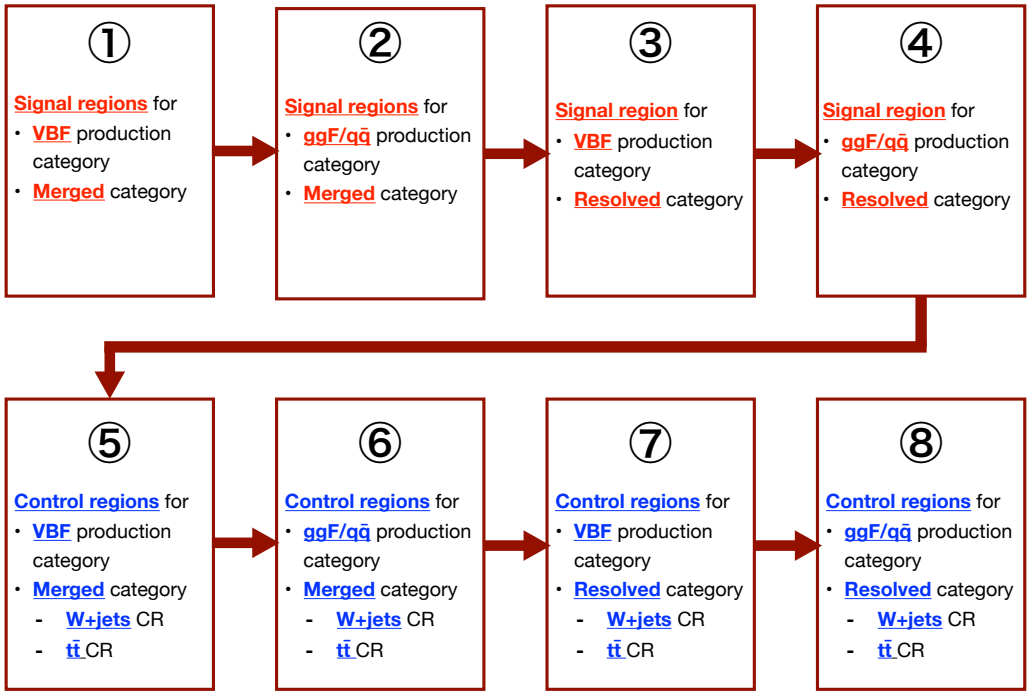


FIGURE 6.11: Flow chart for the definitions of the SR/CRs

Chapter 7

Background estimation

MC simulation is used to model the SM background processes. We estimate the multi-jet background in a data-driven way, while the other backgrounds are estimated by the pure MC predictions. The normalisation factors of the major sources of the background, W+jets and $t\bar{t}$, are estimated in the dedicated control regions, as described in Section 6.3. The normalisation factors of the W+jets and $t\bar{t}$ contributions are treated as free parameters and constrained by fitting in the dedicated control regions. We assign different normalisation factors to the merged and the resolved categories and treated them as uncorrelated. For the merged high-purity (HP) and low-purity (LP) signal regions, we use a set of common background normalisation factors. For a possible difference of the background modelling in the different phase spaces, we employ dedicated normalisation factors for $t\bar{t}$ and W+jets samples for the ggF/$q\bar{q}$ and VBF categories. For the Z+jets, single-t, and dibosons contributions, a Gaussian constraint of 11 %, 11%, and 30 % on the cross section production is applied, based on the SM measurements. The normalisation for these minor backgrounds are treated as correlated between analysis channels.

7.1 Multi-jet background

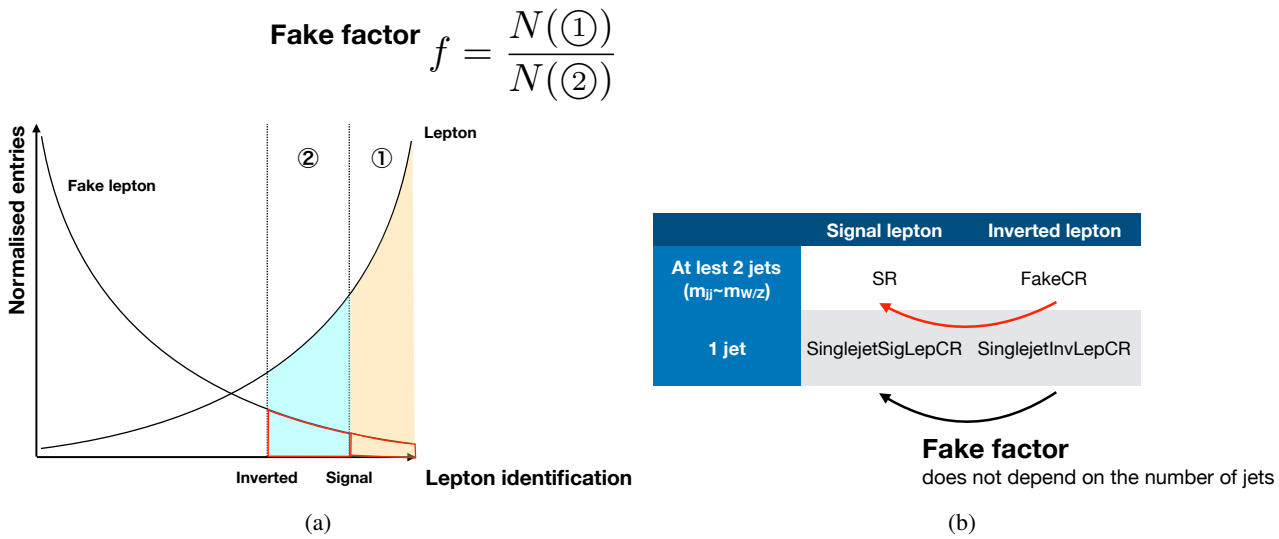
7.1.1 Multi-jet background in the merged category

The multi-jet backgrounds arise from two types of events; (1) QCD events with jets or photon conversions misidentified as leptons, and (2) events with non-prompt leptons from decays of heavy-flavour hadrons. We estimated the multi-jet background in the merged category by a fit to $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of events that satisfy all the signal selection criteria but without any $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. The shape of the multi-jet sample is obtained by applying the same selections with the signal criteria except for the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement and the lepton requirement; the leptons are required to satisfy the veto lepton selection, defined in Section 6.2, but to fail the signal lepton selection. In the fit, the normalisation of the W-jets and multi-jet components are allowed to float, contrary to all the other backgrounds fixed to their predicted cross sections. Following this procedure, the multi-jet background in the merged category is found to be negligible.

7.1.2 Multi-jet background in the resolved category; fake factor method

In the resolved category, we employ a data-driven approach, namely a 'fake factor method', to estimate the multi-jet background. It is important to measure this type of background from data as the rate of misidentification is not accurately modelled in the MC samples. The basic idea of the fake factor method is schematically shown in Figure 7.1(a). We select a control sample of events enriched with the multi-jet background, then derive a fake factor from them to estimate the multi-jet background in the SR. We use two types of lepton identification criteria; (1) signal lepton and (2) inverted lepton. The signal lepton criterion is same with the criterion used in the SR. A looser requirement is imposed on the inverted lepton criterion. The fake factor is calculated as the ratio between the entries which pass the signal and inverted criterion.

In our analysis, we define three control regions; (1) FakeCR, (2) SinglejetSigLepCR, and (3) SinglejetInvLepCR, as shown in Figure 7.1(b). The SinglejetSigLepCR and SinglejetInvLepCR are formed by events that have exactly one small-R jet, while there are at least two small-R jets in the SR and FakeCR. The fake factor does not depend on the number of jets in each event. Hence, we derive the fake factor using these single-jet control regions, to ensure the orthogonality with the SR. We define fake factors (f) as the ratio of the number of events with the



(a) (b)

FIGURE 7.1: (a) Schematic definition of fake factor. To derive the fake factor, we use two types of lepton identification criteria; (1) signal lepton and (2) inverted lepton. By requiring the inverted lepton identification, we select a sample of events enriched with fake leptons. (b) Schematic definitions of the three control regions; (1) FakeCR, (2) SinglejetSigLepCR, and (3) SinglejetInvLepCR. The fake factor does not depend on the number of jets in the events. Thus, we calculate fake factor between the SinglejetSigLepCR and the SinglejetInvLepCR regions, and use the factors to estimate the multi-jet background events in the SR from the FakeCR region.

TABLE 7.1: Requirements for the inverted and the signal electrons (muons) are given in the upper (lower) table. The inverted leptons are used to define the fake factor control regions.

	Inverted electron	Signal electron
Electron identification	*Medium* and !*Tight*	*Tight*

	Inverted muon	Signal muon
Muon track isolation for $p_{\mathrm{T}}(\ell\nu) > 150\,\mathrm{GeV}$	0.06 – 0.15	< 0.06
Muon track isolation for $p_{\mathrm{T}}(\ell\nu) < 150\,\mathrm{GeV}$	0.06 – 0.07	< 0.06

signal lepton selection and with the inverted lepton selection:

$$f = \frac{N_{\mathrm{multijet}}(\mathrm{SinglejetSigLepCR})}{N_{\mathrm{multijet}}(\mathrm{SinglejetInvLepCR})} \quad , \tag{7.1}$$

where N_{multijet}(SinglejetSigLepCR/SinglejetInvLepCR)) is the number of events in the SinglejetSigLepCR/SinglejetInvLepCR region.

The inverted requirements for the leptons are summarised in Table 7.1. We subtract contributions from prompt leptons by applying the inverted requirements for the lepton identification and isolation. This subtraction ensures the higher fraction of fake leptons, whilst being orthogonal to the signal region. Concerning the muon channel with $p_{\mathrm{T}}(\mu\nu) < 150\,\mathrm{GeV}$ where we use an isolated muon trigger, we require the tighter track isolation, i.e. [0.06, 0.07], to avoid trigger bias. At the region of $p_{\mathrm{T}}(\mu\nu) > 150\,\mathrm{GeV}$, on the other hand, we use a $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger thus there is no trigger bias on the muon isolation.

The SinglejetSigLepCR may overlap with the merged SR because the large-R jet can be reconstructed at the same position with the small-R jet. Whereas approximately half of the events with $p_{\mathrm{T}}(j) > 200\,\mathrm{GeV}$ in the SinglejetSigLepCR have a large-R jet, most of the large-R jets have $m(J) < 50\,\mathrm{GeV}$ (note that we apply $m(J) > 50\,\mathrm{GeV}$ in the merged SR/CRs). Thus contamination from the signal events in the merged category into the SinglejetSigLepCR is less than 1 % and it is negligible.

To estimate contribution of the multi-jet background, the contamination from the other SM processes ($t\bar{t}$, W/Z+jets, dibosons, and single-t productions) is subtracted from data using the MC simulated samples. f is highly sensitive to the normalisation of the SM contributions, thus the normalisation of the background shapes in MC is corrected by fitting in the single-jet control regions for $150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$ GeV assuming small contribution from the multi-jet background in this region. This procedure offers a good improvement in the modelling of the multi-jet background. Furthermore, in the electron channel, f is also derived binned in $E_{\mathrm{T}}^{\mathrm{miss}}$. The η bins are chosen based on the detector geometry, while the p_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ bins are optimised such that the number of events in denominator in each bin are almost the same. In the muon channel, the determined binning is:

$|\eta|$: {0, 1.05, 1.5, 2.5} and

$p_{\mathbf{T}}$: {27, 42, 59, 76, 99, inf} GeV,

In the electron channel, it is as follows:

$|\eta|$: {0, 1.37, 1.52, 2.47},

$p_{\mathbf{T}}$: {27, 115, 135, 155, 190} GeV,

$E_{\mathbf{T}}^{\mathbf{miss}}$:
- for $27 < p_{\mathrm{T}} < 115$ GeV : {0, 60, 75, inf} GeV,
- for $115 < p_{\mathrm{T}} < 135$ GeV : {0, 38, 52, inf} GeV,
- for $135 < p_{\mathrm{T}} < 155$ GeV : {0, 26, 43, inf} GeV,
- for $155 < p_{\mathrm{T}} < 190$ GeV : {0, 25, 45, inf} GeV.

Finally, the number of multi-jet background contributions in the signal region N_{multijet} can be estimated as:

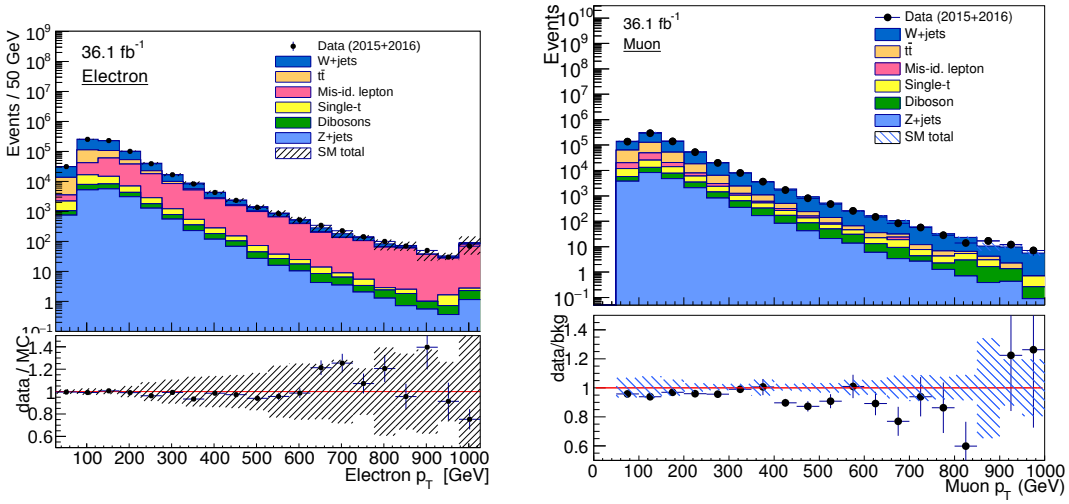
$$N_{\mathrm{multijet}} = f \times N_{\mathrm{multijet}}(\mathrm{FakeCR}) \quad , \tag{7.2}$$

where $N_{\mathrm{multijet}}(\mathrm{FakeCR})$ is the number of the multi-jet background in the FakeCR after the SM background subtraction. In the electron channel, we use 32.9 fb^{-1} data recorded in 2016 in the FakeCR. Therefore, N_{multijet} in total 2015+2016 data is corrected to:

$$N_{\mathrm{multijet}}^{\mathrm{tot}} = f \times N_{\mathrm{multijet}}(\mathrm{FakeCR}) \times \frac{\mathcal{L}}{32.9} \quad , \tag{7.3}$$

where $\mathcal{L}$ is the total integrated luminosity in fb^{-1} in the signal region. In the muon channel, we use the full 2015+2016 data to obtain f thus the scale correction is not necessary.

This estimation is validated in a set of validation region. The validation region (VR) is defined requiring exactly one signal lepton, $30 < E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV and at least one pair of the small-R jets with $p_{\mathrm{T}} > 45$ GeV. The leading jet is required to have $p_{\mathrm{T}} > 60$ GeV. The multi-jet background in the VR is estimated by applying the fake factors to the multi-jet contribution in the fake validation region (FakeVR), requiring the same selections as the VR but requiring inverted lepton. Figure 7.2 shows the lepton p_{T} distribution in the VR. A good agreement between data and prediction is observed both in the electron and the muon channel.



(a) (b)

FIGURE 7.2: Lepton p_{T} distribution in the validation region with $p_{\mathrm{T}}(\ell\nu) > 150\,\mathrm{GeV}$, (a) for the electron channel and (b) for the muon channel. The multi-jet background (Mis-id lepton) is estimated by fake factor method. A good agreement between data and prediction is found both in the electron and muon channel.

Chapter 8

Systematic uncertainties

In this chapter, we summarise the sources of the systematic uncertainties.

8.1 Experimental sources

Each reconstructed object has several sources of uncertainties, each of which is evaluated separately. In this section, we overview the sources of the systematic uncertainties.

8.1.1 Luminosity uncertainties

The uncertainty on the integrated luminosity for the 2015+2016 dataset is 3.2%. The evaluation of the luminosity is performed using several dedicated detectors, and measurements in these detectors are compared to each other to make a precise estimation [71]. The luminosity uncertainty is applied to the MC simulated samples.

8.1.2 Pile-up reweighting uncertainties

MC simulation samples are produced before a given data taking period using a best-guess of the pile-up condition in data. In order to reproduce the pile-up condition in data, we overlay additional pp collisions to MC samples. This procedure is called pile-up reweighting [144]. We consider the uncertainty associated with the pile-up reweighting procedure. A variation in the pile-up reweighting of the MC simulation samples is included to cover the uncertainty on the ratio between the predicted and the measured inelastic cross-sections in the fiducial volume defined by $M_X > 13\,\text{GeV}$, where M_X is the mass of the non-diffractive hadronic system [145].

8.1.3 Trigger uncertainties

The uncertainties on the efficiency of the electron and the muon triggers are evaluated using the tag and probe method, and are applied to the MC simulation samples. The uncertainty from the $E_{\text{T}}^{\text{miss}}$ trigger arises from the estimation on scale factor which consists of two contributions: (1) statistics and (2) discrepancy in the efficiency between MC samples (ttbar and W+jets) [146].

8.1.4 Muons and electrons uncertainties

The following systematic uncertainties are applied to electrons and muons in estimations based on the simulation with the tag-and-probe method using the $Z \to \ell\ell$ events:

- Identification and reconstruction efficiencies: The efficiencies are measured with the tag and probe method using the Z mass peak.
- Isolation efficiency: Scale factor and its uncertainty are derived by the tag and probe method using the Z mass peak as well.
- Energy and Momentum scales: The uncertainties on these values are measured also with Z mass peak.

8.1.5 Missing transverse energy uncertainties

The missing transverse energy is calculated using all the physics objects reconstructed in each event, as described in Section 4.6. Therefore, all the systematic errors on the reconstructed components, e.g. the jet energy scale, result in sources of the systematic uncertainties on $E_{\mathrm{T}}^{\mathrm{miss}}$. There is an additional uncertainty called the "soft term", from the un-associated tracks. The resolution and the scale of the soft term are varied within their errors to evaluate their contribution to the total uncertainty.

8.1.6 Uncertainties on jet energy scale and jet energy resolution

The jet energy scale and resolution of the small-R jets are measured *in-situ* by calculating the response between the MC simulation and the data samples in various bins of the kinematic phase space. We use a globally-reduced parameter configuration, which introduces 21 nuisance parameters in total [147]. These uncertainties are applied not only to the resolved analysis also to the merged analysis because they are used in the calculation of the missing transverse energy. We also consider the uncertainty on the JVT efficiency, which is described in Reference [125].

8.1.7 Large-R jet energy scale and resolution uncertainties

The uncertainty on the mass, p_{T}, and D_2 scale of large-R jet are evaluated by comparing the ratio between the calorimeter-based and track-based measurements using the di-jet events, both in data and MC simulation samples. The combined (TA+calo) mass uncertainty, $\mathrm{Variation}^{\pm}_{\mathrm{comb}}$, is given by:

$$\mathrm{Variation}^{\pm}_{\mathrm{comb}} = \frac{\sigma^2_{\mathrm{TA}}}{\sigma^2_{\mathrm{calo}} + \sigma^2_{\mathrm{TA}}} \times (1 \pm \mathrm{Uncertainty}_{\mathrm{calo}}) + \frac{\sigma^2_{\mathrm{calo}}}{\sigma^2_{\mathrm{calo}} + \sigma^2_{\mathrm{TA}}} \times (1 \pm \mathrm{Uncertainty}_{\mathrm{TA}}) \quad , \tag{8.1}$$

where σ_{calo} and σ_{TA} are the pre-estimated jet mass resolution of the calo- and TA-jet, respectively, as a function of jet p_{T} and η . In our analysis, p_{T} and mass scales are fully correlated while D_2 scale is left uncorrelated ('medium' scenario). There are three possible scenarios for the correlation of the three parameters:

- strong: p_{T}, mass, and D_2 scales are fully correlated. Tracking and modelling components are combined.
- medium: p_{T} and mass scales are fully correlated while D_2 scale is left uncorrelated. Tracking and modelling components are left independent.
- weak: p_{T} and mass scales are left uncorrelated. Tracking and modelling components are left independent.

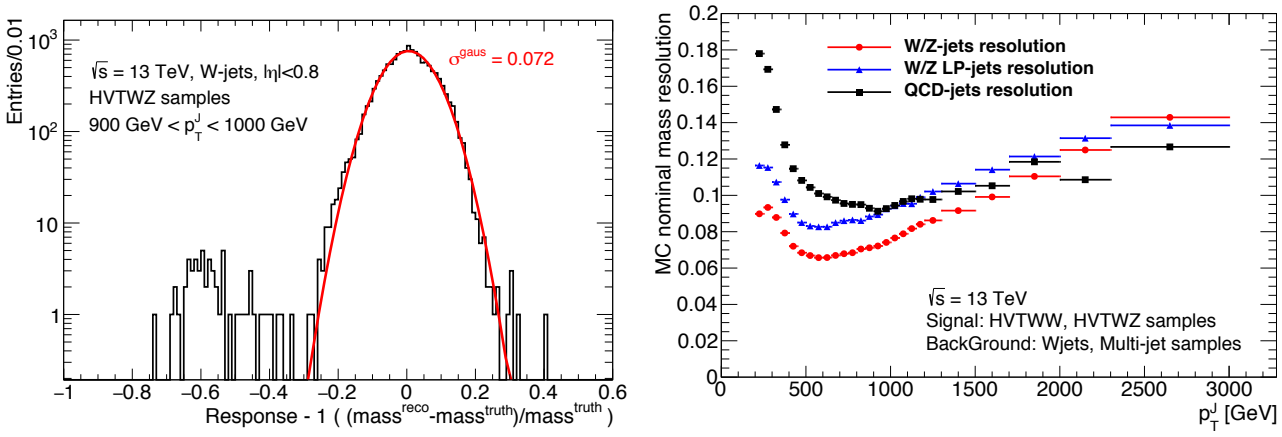
We try each of the three scenarios in the fit and find that the obtained results (e.g. upper limit on the cross section) are not different among the three scenarios. Thus, we adopt the medium scenario in this analysis for the consistency with the other analysis channels.

The large-R jet resolution uncertainties are applied as follows:

- $\mathrm{p_T}$ resolution uncertainty: '*absolute*' 2%,
- Mass resolution uncertainty: '*relative*' 20%,
- Any substructure resolution uncertainties: '*relative*' 15%.

The 'absolute' uncertainty is applied by smearing p_{T} with a Gaussian function with $\sigma = p_{\mathrm{T}} \times 2\%$. The goal of the 'relative' resolution uncertainties is to smear the reconstructed observable (mass, D_2) by a Gaussian smearing, such that the nominal resolution ($\sigma_{\mathrm{nominal}}$) is increased by 20 % or 15 % ($\sigma_{\mathrm{new}} = 1.2/1.15 \cdot \sigma_{\mathrm{nominal}}$) as appropriate. Following Equation 8.2, to increase the resolution by 20/15 %, we need to smear the original distribution (mass or D_2) by a Gaussian with a width of $0.66/0.57\sigma_{\mathrm{nominal}}$:

$$\begin{aligned} \sigma^2_{\mathrm{new}} &= \sigma^2_{\mathrm{nominal}} + (x \cdot \sigma^2_{\mathrm{nominal}}) \quad , \\ (1.2/1.15 \cdot \sigma_{\mathrm{nominal}})^2 &= (1 + x^2)\sigma^2_{\mathrm{nominal}} \quad , \\ 1.44/1.32 &= 1 + x^2 \quad , \\ x &= 0.66/0.57 \quad . \end{aligned} \tag{8.2}$$



(a) (b)

FIGURE 8.1: (a) Example distribution of mass response for W-jets in $900 < p_{\mathrm{T}}{}^{\mathrm{J}} < 1000$ GeV. (b) The evaluated nominal mass resolution binned in p_{T}. The nominal resolution is obtained by applying a Gaussian fit to the mass response distribution.

It means that the MC nominal resolution should be evaluated by determining the width of the response distribution of a given observable for the jets of interest. This is done by fitting the response distribution with a Gaussian function and taking the width of this function as an estimate of the resolution.

Figure 8.1(a) shows an example of the mass response distribution for W-jets. The response follows a Gaussian distribution, therefore the nominal resolution in MC is determined as the sigma of the Gaussian. The nominal resolution in MC is evaluated separately for W/Z-jets, W/Z 'low-purity jets', and QCD-jets binned in p_{T} and η; here the large-R low-purity jets fail either the 50% D_2 cut or the 50% mass cut. There is little difference between the resolution of W-jets and Z-jets, and that means the η-dependence is small. Thus the results of W- and Z-jets are merged. The η bins are also merged. Figure 8.1(b) shows the evaluated resolution. Following Equation 8.2, the smearing value which can increase the resolution by 20 % can be calculated from the MC nominal resolution as a factor of 0.66 for the nominal resolution. To be on the conservative side, 2 % is added as a safety margin:

$$\text{smearing value} = \sigma_{\text{nominal}} \times 0.66 + \text{safety margin}(2\%) \quad . \tag{8.3}$$

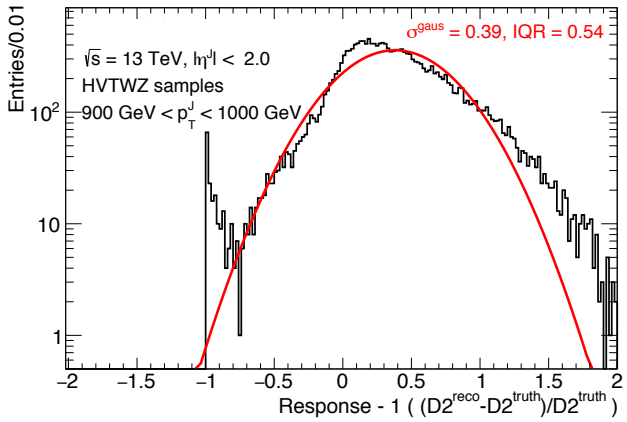
The nominal resolution of D_2 is evaluated in a similar way. The D_2 response distribution does not follow a Gaussian[1], however as a rough estimation, the fitted gaussian sigma is adopted as the MC nominal resolution. Figure 8.2(a) shows an example of the D_2 response distribution for the W/Z-jets. As shown in the figure, the result of the fitting with a Gaussian is not very good. Furthermore, there is strong dependence of the resolution on the D_2 value it self, as shown in the Figure 8.2(b) and 8.2(c). Following Equation 8.2, the smearing value which can increase the resolution by 15 % can be calculated from the MC nominal resolution as a factor of 0.57 for the nominal resolution. To be on the conservative side, 2 % is added as a safety margin.

$$\text{smearing value} = \sigma_{\text{nominal}} \times 0.57 + \text{safety margin}(2\%) \quad . \tag{8.4}$$

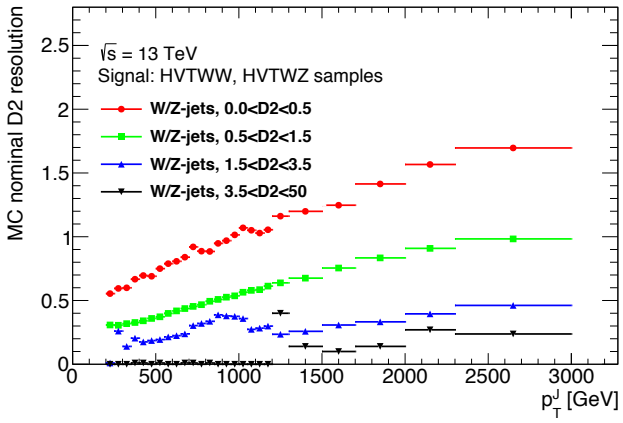
As a result of this study, we apply very conservative uncertainty on the D_2 resolution. We observe strong over-constraint on the large-R jet D_2 resolution uncertainty in the post-fit uncertainties as shown in Chapter 10, however it is as expected. [2]

[1] D_2^{truth} is calculated using the truth particles, instead of the energy deposit in the calorimeter.

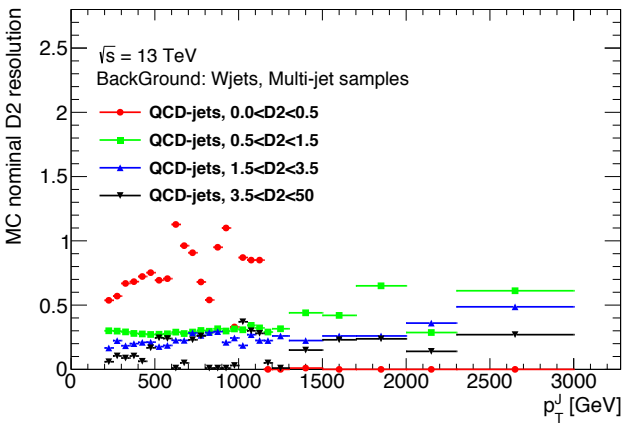
[2] To avoid such over-estimation of the uncertainty, we are developing a different method to evaluate the resolution uncertainty of the large-R jet D_2. Please refer to Reference [148].



(a)



(b)



(c)

FIGURE 8.2: (a) Example of D_2 response distribution for W/Z-jets in $900 < p_T{}^J < 1000$ GeV, (b) the evaluated nominal mass resolution of W/Z-jets binned in p_T, and (c) the evaluated nominal mass resolution of QCD-jets binned in p_T. The nominal resolution is obtained by applying a Gaussian fit to the D_2 response distribution.

8.1.8 *b*-tagging uncertainties

The systematic uncertainties associated to the *b*-tagging are considered [126]. They are evaluated as uncertainties on the scaling factor to take into account for possible disagreement on the *b*-tag efficiency between the data and the MC samples. Separate scale factors and corresponding systematic uncertainties are provided for *b*-, *c*-, and light-flavour-induced jets, based on several measurements. We assign a scale factor for the *c*-induced jets to τ-induced jets and an additional uncertainty is considered to account for the *c* to τ extrapolation. The number of systematic variations are reduced to 3, 4, and 5 eigenvectors for the *b*-, *c*-, and light-flavour-induced jets, respectively, considering correlation between the variations. We introduce an additional uncertainty due to the extrapolation for jets with p_{T} beyond the kinematic reach of the calibration data.

8.2 Signal uncertainties

8.2.1 Uncertainties on initial/final state radiations

In order to evaluate the uncertainties from the initial state radiation (ISR) and final state radiation (FSR), we introduce the following five sets of eigenvectors [107]:

- Var1 : mainly for underlying event effects,
- Var2 : mainly for jet substructure effects,
- Var3 a/b/c : for the different aspects of extra jet production.

The effects of the ISR/FSR variations change some distributions e.g. the D_2 distribution, as shown in Figure 8.3. The differences of the signal acceptance in the ISR/FSR variation samples from the nominal sample are estimated by using the truth MC samples. These five variations are summed in quadrature to get the full coverage of the uncertainty. Figure 8.4 shows the evaluated uncertainty on each of the HVTWZ and the bulk graviton samples. Each result is parametrized using a cubic (quadratic) function in the Merged (Resolved) category. The uncertainties are evaluated to be 4 % except for the low-mass samples (mass < 400 GeV). It becomes large at $m > 1$ TeV in the resolved analysis, but at that region signal acceptance is very low as shown in Figure 6.9 and 6.10. Thus there is no problem to assign a conservative uncertainty at that region to take into account the statistical fluctuation of the MC samples.

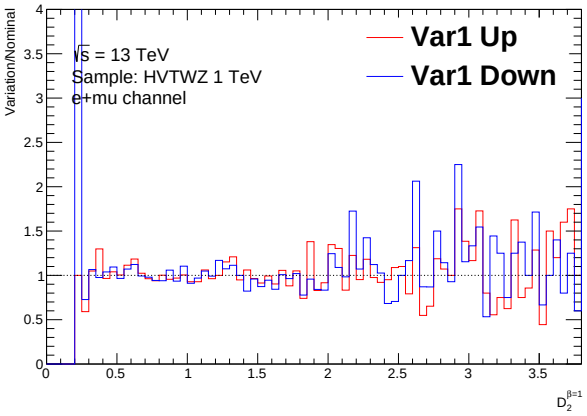
8.2.2 PDF uncertainties

The PDF uncertainty on the signal yield is evaluated by comparing the nominal PDF set with PDF variation sets (MMHT-2014 [75] and CT14 [149]). The PDF uncertainty is evaluated as the combination of the 68% uncertainty band for each of the PDF sets, on the signal acceptance relative to the nominal:

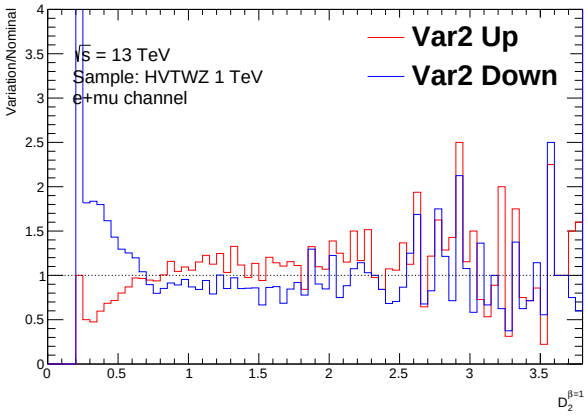
$$\frac{(\varepsilon \times A)_{\mathrm{var}}}{(\varepsilon \times A)_{\mathrm{nom}}} = \frac{\sum\limits_{\mathrm{SR}} w_{\mathrm{var}}}{\sum\limits_{\mathrm{all}} w_{\mathrm{var}}} \Bigg/ \frac{\sum\limits_{\mathrm{SR}} w_{\mathrm{nom}}}{\sum\limits_{\mathrm{all}} w_{\mathrm{nom}}} \quad , \tag{8.5}$$

where w_{var} and w_{nom} are the event weights of the variation and the nominal PDF choice, respectively; the sums are taken over the events in the signal region and all the events [150].

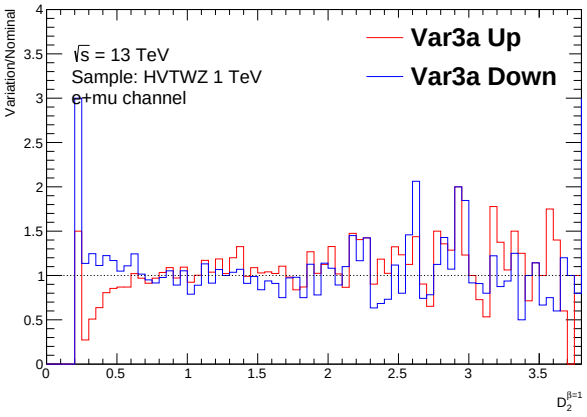
Figures 8.5 and 8.6 show the uncertainties on the combined signal acceptance relative to the nominal PDF for the ggF/$q\bar{q}$ and the VBF production signals. The uncertainties are evaluated to be less than 2% for most of the signal samples. To be on the conservative side, flat 0.5 % error is assigned for the heavy higgs and the bulk Graviton samples, while flat 2 % error is assigned for the HVT samples. For the signals produced via the VBF process, flat 1 % error is assigned for the heavy higgs samples and flat 2 % is assigned for the HVT samples.



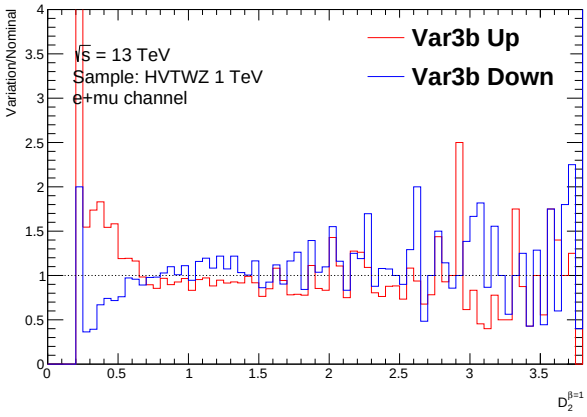
(a) Var1



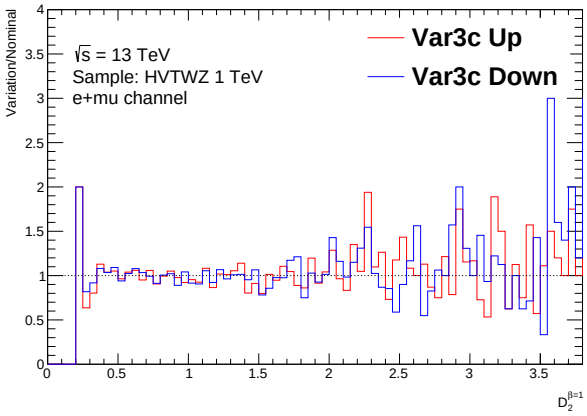
(b) Var2



(c) Var3a

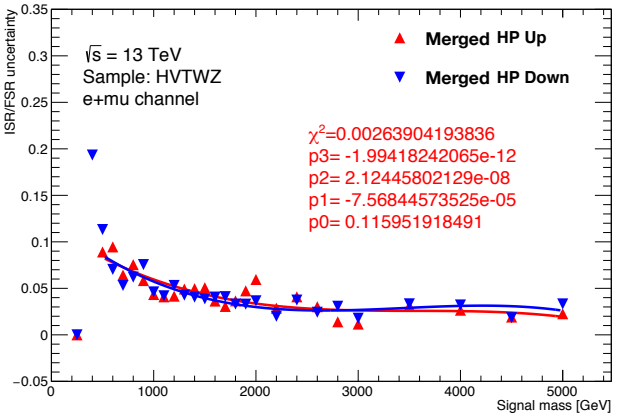


(d) Var3b

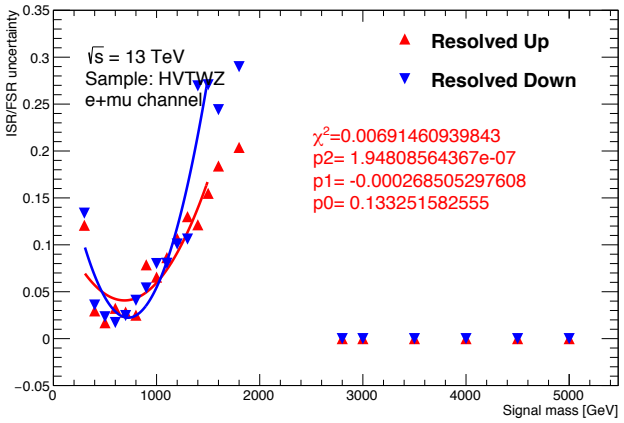


(e) Var3c

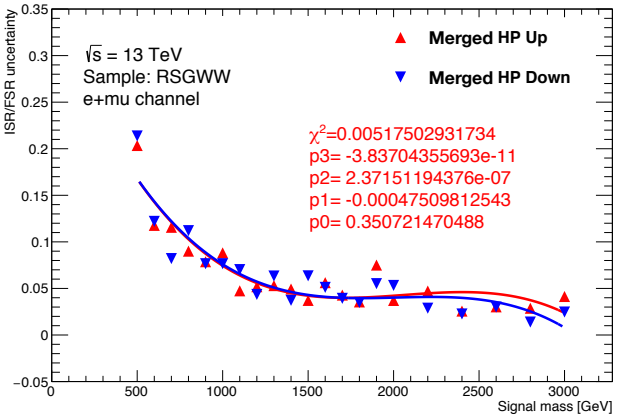
FIGURE 8.3: D_2 distributions with the ISR/FSR variation in the HVT samples. These five variations are summed in quadrature to get the full coverage of the uncertainty.



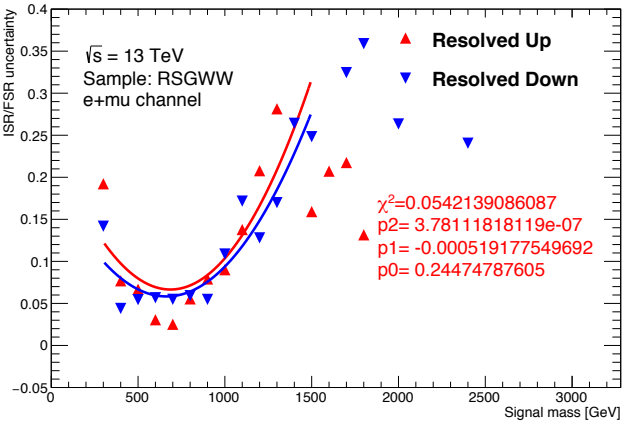
(a) HVT Merged



(b) HVT Resolved

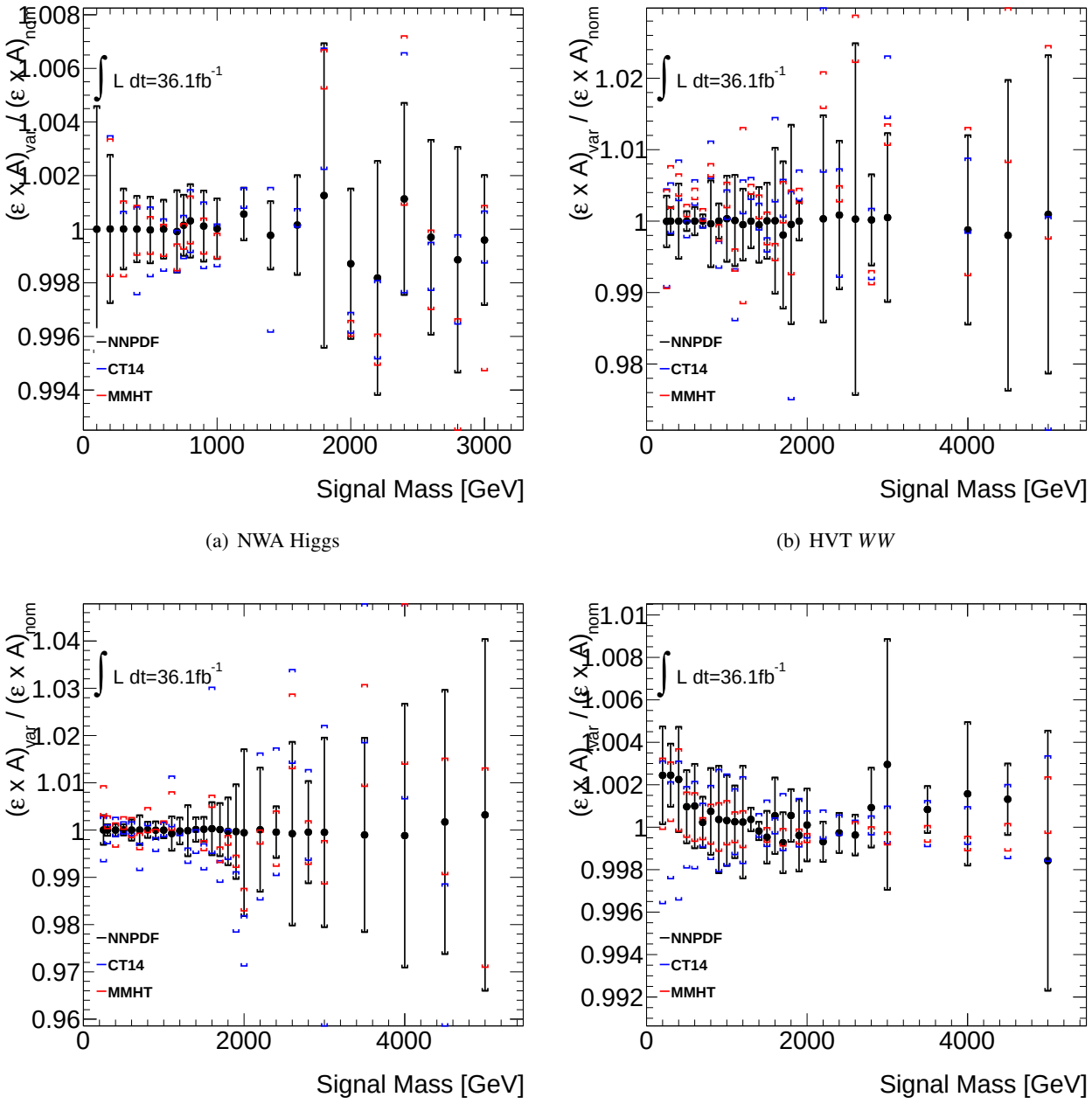


(c) bulk Graviton Merged



(d) bulk Graviton Resolved

FIGURE 8.4: ISR/FSR uncertainties evaluated for the HVT (bulk Graviton) samples in the Merged (Resolved) regions. Each of the results is parametrized using a cubic (quadratic) function in the Merged (Resolved) category. The uncertainties are evaluated to be 4 % except for the low-mass samples (mass < 400 GeV).



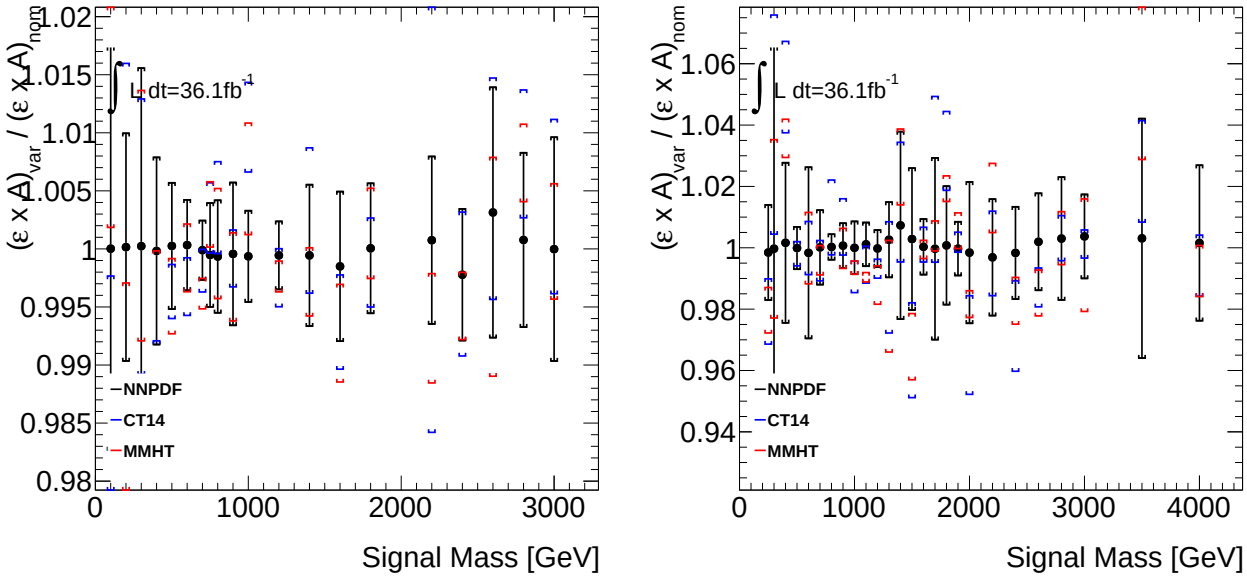
(a) NWA Higgs

(b) HVT WW

(c) HVT WZ

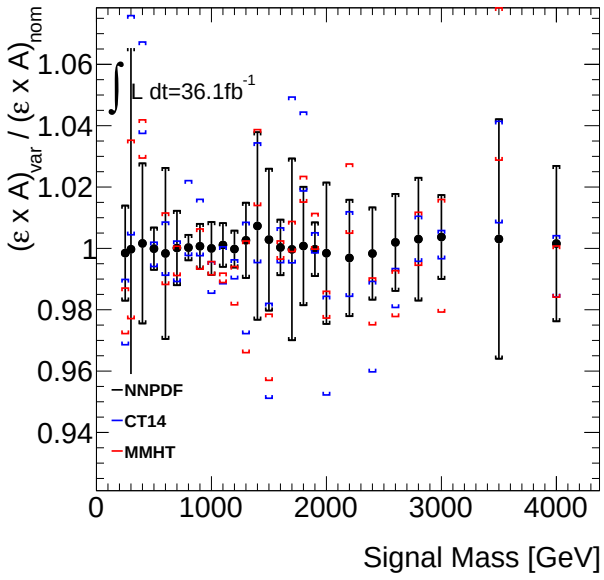
(d) bulk Graviton

FIGURE 8.5: PDF uncertainties in the ggF$q\bar{q}$ category from the NNPDF3.0 (black), MMHT(blue), and the CT14(red) on the relative signal acceptance for each mass point produced via the ggF/$q\bar{q}$ process. The envelop of these errors is chosen as the PDF uncertainty.



(a) NWA Higgs

(b) HVT *WW*



(c) HVT *WZ*

FIGURE 8.6: PDF uncertainties in the VBF category from the NNPDF3.0 (black), MMHT(blue), and the CT14(red) on the relative signal acceptance for each mass point produced via the VBF process. The envelop of these errors is chosen as the PDF uncertainty.

Chapter 9

Statistical analysis

In this chapter, we overview the quantitative strategy used to compare the recorded data to the SM predictions, and to set limits on the BSM theoretical models. The invariant mass, $m(WV)$, is used as a discriminant in this analysis, with an optimised binning that is explained in detail in this chapter.

9.1 Maximum likelihood fit

We perform a binned maximum-likelihood (ML) fit to the m(WV) distributions using templates derived from the MC simulation and the data-driven estimates using a standard package, RooStats [151]. The fit is performed on 6 (7) components in the Merged (Resolved) categories : signal, W+jets, $t\bar{t}$, single-t, Z+jets, and dibosons (+multijet). The electron and muon channels are merged in the fit procedure.

We define the m(WV) binning considering the expected resolution of the resonance peak. We use a set of signal MC samples (HVT Z' for the merged and NWA $H \rightarrow WW$ for the resolved) to estimate the width of the resonance; the width of the peak is shown as a function of the resonance mass and it is fitted with a polynomial function as shown in Figure 9.1. The obtained polynomial function is used to define the optimal set of variable bin widths. In the high mass region, the binning is coarser to avoid bins with too low statistics. According to the signal efficiency shown in Figure 6.9 and 6.10, we use the region of $m(\ell\nu jj) < 1.5$ TeV in the resolved channel. In the VBF category, the binning is

[500,575,660,755,860,975,1100,1235,1380,1535,1700,1875,2060,2255,2460,2675,2900,3135,3380,3800,15000]

in the unit of GeV, for the merged analysis and

[300,360,420,500,575,660,755,860,975,1100,1500]

in the unit of GeV for the resolved analysis in the ggF/$q\bar{q}$ category.

Due to the lack of MC statistics for the background samples, we use a coarser binning at the high-mass region in the VBF category. The determined binning for the merged analysis is:

[500,575,660,755,860,975,1100,1235,1380,1535,1875,15000]

in the unit of GeV, and

[300,360,420,500,575,660,755,860,975,1100,1500],

in the unit of GeV for the resolved analysis.

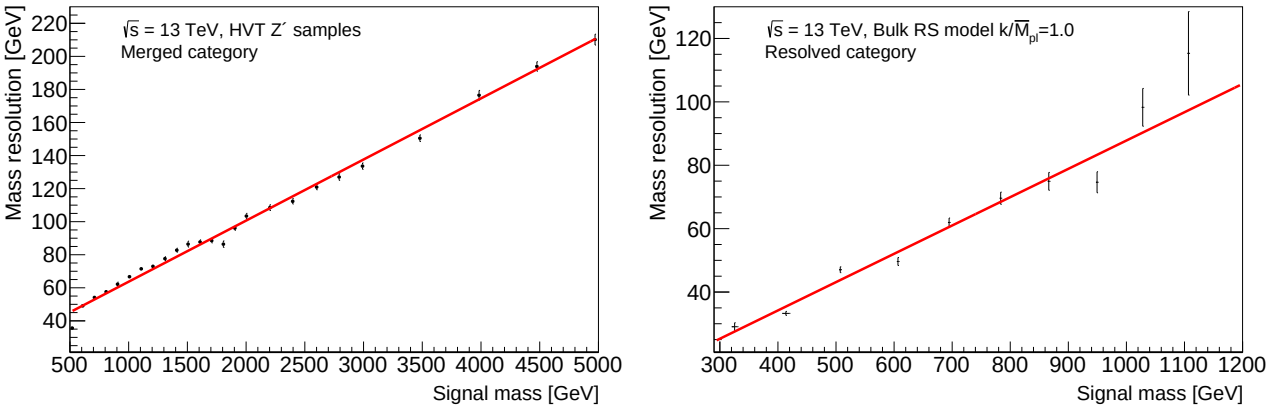
9.1.1 Combined fit

We perform a combined fit on the events in the signal regions, the top control regions, and the W+jets control regions. The likelihood used in this analysis is written as:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_k \prod_j \prod_i P(N^j_{ki} | \mu s^j_{ki} + B^j_{ki}) \times \prod_l \text{Nuis}(\tilde{\theta}_l) \quad , \tag{9.1}$$

where B^j_{ki} is the background expectation described as:

$$B^j_{ki} = \mu_{t\bar{t},k} b^j_{t\bar{t},ki} + \mu_{W,k} b^j_{W,ki} + b^j_{\text{others},ki} \quad . \tag{9.2}$$



(a) (b)

FIGURE 9.1: Width of the resonance peak as a function of the resonance mass in the ggF/$q\bar{q}$ category (a) for the merged and (b) the resolved analysis.

	ggF Higgs HVT RS G^*	VBF Higgs VBF HVT
Merged HP ggF SR/CR	✓	
Merged LP ggF SR/CR	✓	
Resolved ggF SR/CR	✓	
Merged HP VBF SR/CR		✓
Merged LP VBF SR/CR		✓
Resolved VBF SR/CR		✓

TABLE 9.1: Analysis channels involved in each interpretation.

Here, i is the bin number in the m(WV) distributions; j is the index of the region in the channel k[1]; k is the index of the search channels[2]; $P(a|b)$ is the Poisson probability to observe a if b is expected; μ is the signal strength treated as the common parameter among all the regions; $\mu_{t\bar{t},k}$ and $\mu_{W,k}$ are the normalisation factors for the $t\bar{t}$ and W+jets contributions in a given analysis channel k; N^j_{ki} is the number of observed data events in the bin i in the region j of the channel k; s^j_{ki} and $b^j_{x,ki}$ are the expected number of yields for the signal and background component x, respectively, in the bin i in the region j of the channel k, where x is $t\bar{t}$, W+jets, and the other background components.

The regions used in each interpretation are summarised in Table 9.1. The VBF categories are not involved for the ggF/$q\bar{q}$ interpretations since no sensitivity gain is found with adding them (acceptance times efficiency is almost 0). It is possible to improve the sensitivity by only a few % by adding the ggF/$q\bar{q}$ categories to the VBF interpretations Therefore, for the simplicity, we do not use the ggF/$q\bar{q}$ regions in the VBF interpretations neither.

The signal and background expectations are obtained as functions of the nuisance parameters $\boldsymbol{\theta}$. For each of the systematic uncertainties, we define a pair of "up" and "down" histograms made by varying a parameter by $\pm 1\sigma$. The nuisance parameter θ_l describes a possible variation in the shape and normalisation of the histograms in the systematic uncertainty l. In each channel k, region j, and bin i, the expected signal and background yields are described as:

$$\mu s + B = (\mu s + B)^0 \prod_l \left(1 + \frac{\Delta_l}{(\mu s + B)^0}\right) \quad , \tag{9.3}$$

[1] $j = 0$: signal region (SR), $j = 1$: top control region (TR) and $j = 2$: W+jets control region (WR)

[2] $k = 0$: Merged High-Purity ggF/$q\bar{q}$; $k = 1$: Merged Low-Purity ggF/$q\bar{q}$; $k = 2$: Merged High-Purity VBF; $k = 3$: Merged Low-Purity VBF, $k = 4$: Resolved ggF/$q\bar{q}$; $k = 5$: Resolved VBF

where

$$\Delta_l = \begin{cases} \theta_l \left\{ (\mu s + B)^+ - (\mu s + B)^0 \right\}, \theta \geq 0 \\ \theta_l \left\{ (\mu s + B)^0 - (\mu s + B)^- \right\}, \theta < 0 \end{cases} , \tag{9.4}$$

where $(\mu s + B)^0$, $(\mu s + B)^+$, and $(\mu s + B)^-$ are the entries in the nominal, up, and down histograms in the bin i. A Gaussian constraint with mean 0 and width 1 is applied to each of the θ_l, which is described by 'Nuis' in Equation 9.1. All the nuisance parameters are treated as correlated in all the regions in every single analysis channel.

The template histograms are obtained based on simulation samples with finite-statistics (or a control-region with finite-statistics in the case of the multi-jet template), thus we also introduce nuisance parameters to account for the bin-by-bin statistical uncertainties on the template histograms. We use the Barlow-Beeston "lite" method [152], in which we introduce a nuisance parameter for each bin in the histogram and in each channel; the nuisance parameter represents the statistical errors in that bin from each of the components added in quadrature.

9.1.2 Smoothing of the systematic uncertainties

To avoid the effect of statistical fluctuation, we apply a smoothing procedure to the template histograms. For each bin i, the entries of the histograms are set to:

$$\begin{aligned} (\mu s + B)_i^+ &= (\mu s + B)_i^0 \times (1 + \lambda_i^+) \quad , \\ (\mu s + B)_i^- &= (\mu s + B)_i^0 \times (1 + \lambda_i^-) \quad , \end{aligned} \tag{9.5}$$

where

$$\begin{aligned} \lambda_i^+ &= \frac{\left\{ (\mu s + B)_{i-1}^+ - (\mu s + B)_{i-1}^0 \right\} + 2 \left\{ (\mu s + B)_i^+ - (\mu s + B)_i^0 \right\} + \left\{ (\mu s + B)_{i+1}^+ - (\mu s + B)_{i+1}^0 \right\}}{4} \quad , \\ \lambda_i^- &= \frac{\left\{ (\mu s + B)_{i-1}^- - (\mu s + B)_{i-1}^0 \right\} + 2 \left\{ (\mu s + B)_i^- - (\mu s + B)_i^0 \right\} + \left\{ (\mu s + B)_{i+1}^- - (\mu s + B)_{i+1}^0 \right\}}{4} \quad . \end{aligned} \tag{9.6}$$

9.2 Signal significance and Upper limits on cross section

In order to quantify the level of agreement, a 'p-value' is defined from the probability distribution function (PDF) of the test statistic ($\tilde{q_\mu}$), $f(\tilde{q_\mu}|\mu)$:

$$p_\mu = \int_{\tilde{q_\mu},\mathrm{obs}}^{\infty} f(\tilde{q_\mu}|\mu) d\tilde{q_\mu} \quad . \tag{9.7}$$

Asymptotic distributions are taken from Reference [153] and used for the signal mass points below 1.6 TeV (1.0 TeV) for fits to the regions with the ggF (VBF) selection. In this method, the test statistic is computed as:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \\ -2 \ln \frac{L(0, \hat{\hat{\theta}}(0))}{L(\hat{\mu}, \theta(\hat{\mu}))} & \hat{\mu} < 0 \end{cases} , \tag{9.8}$$

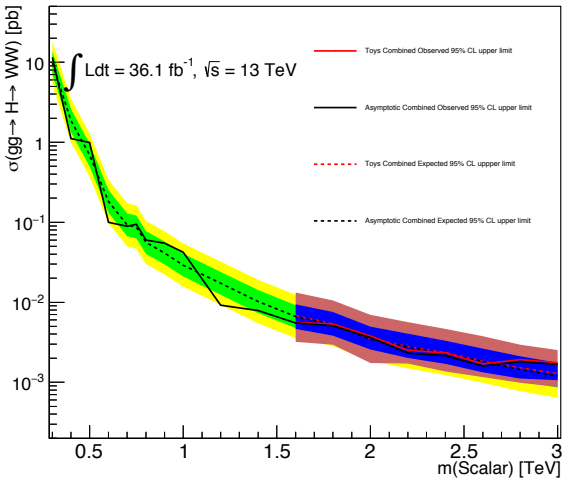
where L is the likelihood function, $\hat{\mu}$ and $\hat{\theta}$ are the best fit values for μ and θ with all parameters floating; and $\hat{\hat{\theta}}(\mu)$ is the best fit value for θ for a fixed value of μ. The statistic represents the compatibility of the data with the μ hypothesis using the ratio of likelihood for the case of a floating μ (denominator), and the case where it is fixed at the hypothesis value (nominator). As for the case of q_0, a one sided prescription is used, assigning $\tilde{q}_\mu = 0$ if the fitted value $\hat{\mu}$ is above the hypothesis. Finally, if $\hat{\mu} < 0$, the case $\mu = 0$ is used instead to avoid technical issues from negative PDFs.

We evaluate upper limits on the signal strength using a modified frequentist approach (CL_s) [154]. A CL_{s+b} value is computed using an asymptotic method [153]. The modified CL_s exclusion is then computed as:

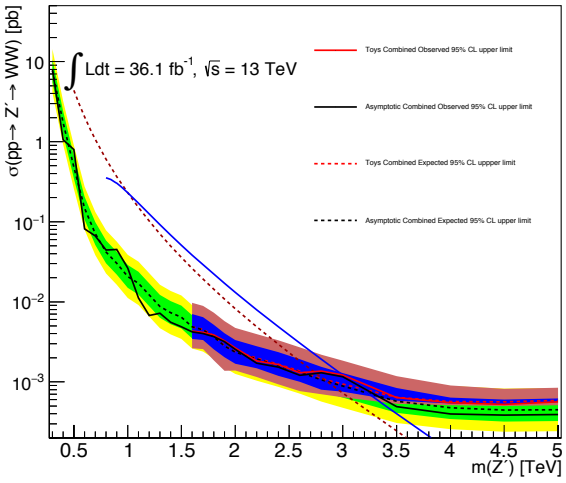
$$CL_s = \frac{CL_{s+b}}{CL_b} \quad , \tag{9.9}$$

where the CL_b value is obtained in the same way as CL_{s+b} but in the background only hypothesis. Limits at 95% confidence level on the value of the signal strength μ are then computed by scanning values of the μ hypothesis, computing the CL_s exclusions and identifying the μ_{up} for which this value equals 0.05.

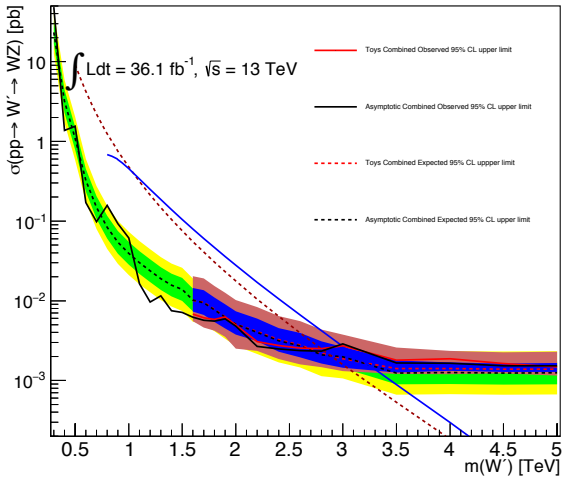
For higher signal masses, the PDFs are estimated with an ensemble of 50,000 generated pseudo-experiments using toy MC methods. We confirm that the asymptotic and pseudo-experiment limits agree well near the boundary between the two regions as shown in Figure 9.2 and 9.3. There is low statistics in the high mass region, thus we need to use a poisson function to evaluate the distribution, while in an asymptotic approximation, we use a gaussian function to evaluate it. In the low mass region, there is enough statistics thus we can use the asymptotic approximation following the central limit theorem.



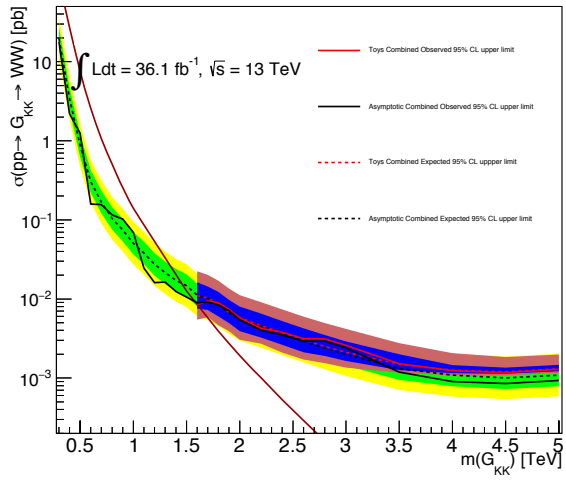
(a)



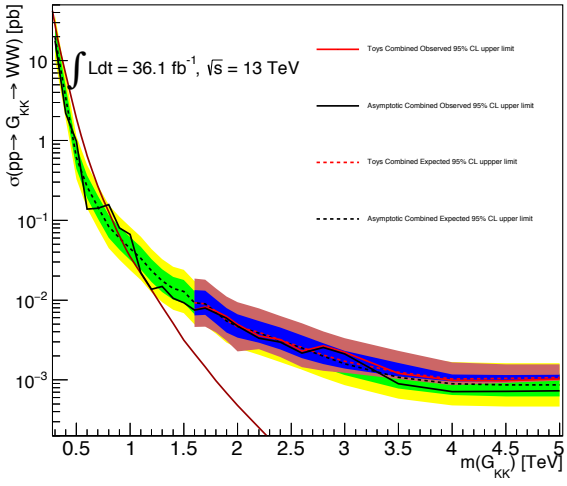
(b)



(c)

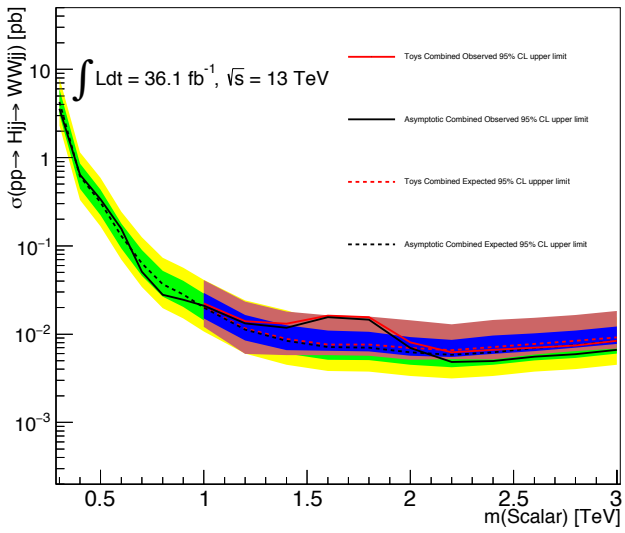


(d)

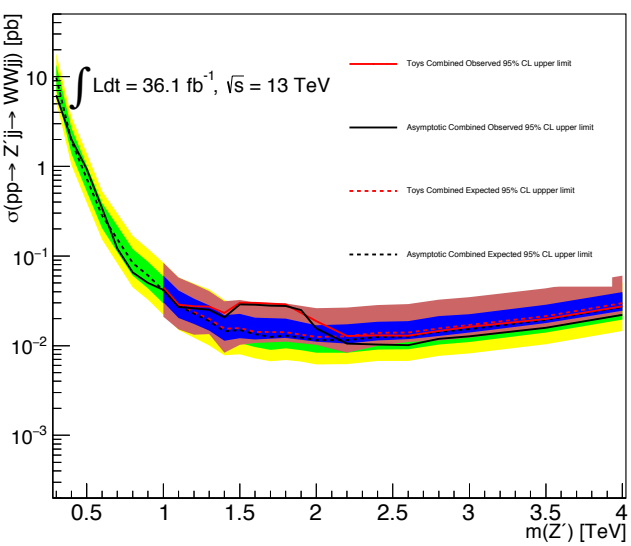


(e)

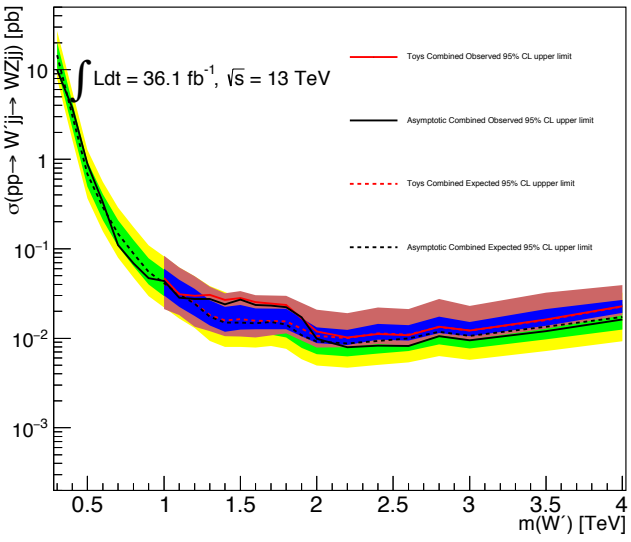
FIGURE 9.2: Limit comparison between the asymptotic and the toy limits for the ggF/$q\bar{q}$ signal samples; (a) scalar $H \to WW$, (b) HVT Z', (c) HVT W', (d) G_{KK} $k/\bar{M}_{Pl} = 1.0$, and (e) G_{KK} $k/\bar{M}_{Pl} = 0.5$. The observed and the expected upper limits on the cross-section at the 95% confidence level in the ggF/qq̄ category are presented as a function of the resonance mass. The red (black) solid/broken lines show the observed/expected upper limits estimated by the asymptotic (toy MC) method. These two types of the upper limits agree well at ~ 1.6 TeV.



(a)



(b)



(c)

FIGURE 9.3: Limit comparison between the asymptotic and the toy limits for the VBF signal samples; (a) scalar $H \to WW$, (b) HVT Z', and (c) HVT W'. The observed and the expected upper limits on the cross-section at the 95% confidence level in the VBF category are presented as a function of the resonance mass. The red (black) solid/broken lines show the observed/expected upper limits estimated by the asymptotic (toy MC) method. These two types of the upper limits agree well at $\sim$ 1.0 TeV.

Chapter 10

Results

10.1 Summary of the analysis flow

In this section, we briefly summarise the analysis flow. We categorise the events depending on the production mechanism and the event topology.

- Production mechanism:
  - Merged category;
  - Resolved category;
- Event topology:
  - VBF category;
  - ggF/$q\bar{q}$ category;

After the event selections, the major contributions from the SM background are top quark pair production ($t\bar{t}$) and W boson production associated with hadron jets (W+jets). These contributions are normalised in a set of dedicated control regions (CRs). This flow and the corresponding figures and tables are summarised in Figure 10.1.

10.2 Post-fit distributions

We combined the merged and the resolved analyses forming a product of the likelihoods of the BSM signals for each of the analyses as explained in Chapter 7, in the mass range from 500 GeV to 1.4 TeV. On the other hand, resolved-only and merged-only fits are performed in the regions of $m < 500$ GeV and $m > 1.4$ TeV, respectively. The systematic uncertainties are taken into account as constrained nuisance parameters with Gaussian or log-normal distributions. The signal strength, μ, and other nuisance parameters that appear in all analyses are treated as correlated. The only exceptions are the $t\bar{t}$ and W+jets normalisation, which are only correlated between the merged High-Purity and Low-Purity regions. After combining these analyses, we performed the profile likelihood fits.

The normalisation factors in the best fit for the background estimation are summarised in Table 10.1. The expected and observed post-fit background yields are summarised in Tables 10.2, 10.3, 10.4, and 10.5 for the WW channel in the ggF/$q\bar{q}$ category, the WZ channel in the ggF/$q\bar{q}$ category, the WW channel in the VBF category,

TABLE 10.1: Normalisation factors for the main background sources, namely W+jets and $t\bar{t}$, in the VBF and ggF/$q\bar{q}$ categories. The fit procedure is described in Section9.1. The factors defined as the ratio of the number of fitted events to the number of predicted events in simulation, The quoted errors incorporate statistical and systematic uncertainties [38].

Category	Signal Region	WW Selection		WZ Selection	
		W+jets	$t\bar{t}$	W+jets	$t\bar{t}$
VBF	Merged	0.89 ± 0.18	1.21 ± 0.18	0.84 ± 0.16	1.10 ± 0.17
	Resolved	1.13 ± 0.25	1.22 ± 0.18	1.08 ± 0.25	1.21 ± 0.17
ggF/$q\bar{q}$	Merged	0.95 ± 0.06	1.03 ± 0.06	0.97 ± 0.06	1.00 ± 0.06
	Resolved	1.06 ± 0.08	1.02 ± 0.05	1.06 ± 0.08	1.00 ± 0.05

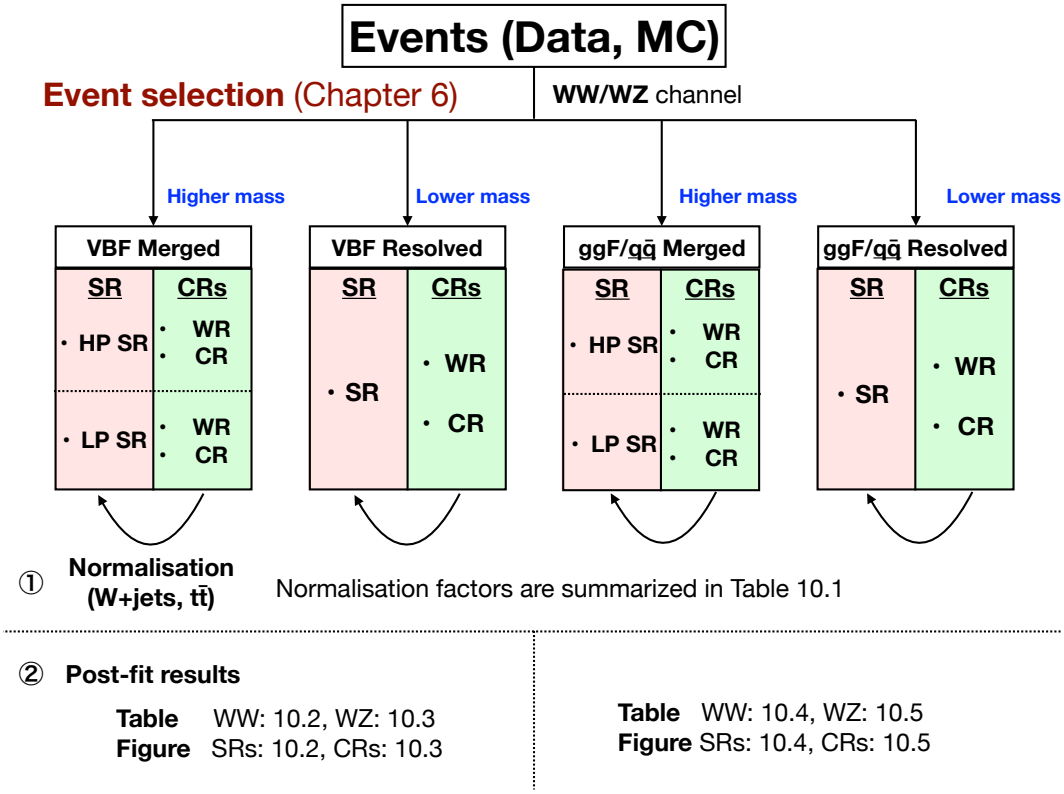


FIGURE 10.1: Brief schematic summary of the analysis flow.

and the WZ channel in the VBF category, respectively. Figures 10.2, 10.3, 10.4, and 10.5 show the post-fit m_{WV} distributions for the ggF/$q\bar{q}$ SRs, the ggF/$q\bar{q}$ CRs, the VBF SRs, and the VBF CRs, respectively. The post-fit distribution for D_2 is presented in Figure 10.6, in which the High- and Low-Purity signal regions and control regions are combined. The fit results describe the observed data very well, and no significant excesses are observed with regard to the SM expectations.

10.3 Testing of signal hypotheses

The local p_0 values are shown in the Figures 10.7 and 10.8 for the ggF/$q\bar{q}$ and VBF category, respectively. The largest local excess is observed at $m(Z') = 1.7$ TeV with the VBF production with significance of approximately 2.6 σ for the signal. In the post-fit plots, Figure 10.4, there are three events observed in the bin at 1.7 TeV while less than one event expected.

10.4 Expected and Observed upper limits

The expected and observed upper limits on the cross section of new particles decaying to WW/WZ are calculated by combining the merged high- and low-purity signal regions with the corresponding resolved region. For resonance masses below 1.0 (1.6) TeV in the VBF (ggF/$q\bar{q}$) category, the upper limits are evaluated using the CL_s method [154], in the asymptotic approximation, at the 95% confidence level (CL). For higher masses, the small number of expected events makes the asymptotic approximation imprecise (it follows the Poisson distribution instead of the Gaussian distribution) and the limits are calculated using pseudo-experiments as discussed in Section 9.2. The calculated upper limits for all benchmark signal models are shown in Figures 10.9 and 10.10, and 10.11, for the ggF/$q\bar{q}$ and the VBF category, respectively. The predicted signal cross-sections are overlaid, except for two types of signals; (1) the scalar signals, which are used for feasibility studies of searches for extended Higgs sectors, (2) the signals produced via the VBF mechanism, to which we can not set any exclude limits. The interpretation in the VBF (ggF/$q\bar{q}$) category assumes that there is no signal leakage from the ggF/$q\bar{q}$ (VBF) processes.

TABLE 10.2: Expected and observed number of events in the signal and control regions for the WW channel, evaluated after a background-only fit to the data in the ggF/$q\bar{q}$ category. The uncertainty on estimation of the total background can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources [38].

	WW Signal region	W+jets Control region	Top Control region
	Merged, High Purity		
W+jets	3120 ± 170	6850 ± 210	540 ± 60
$t\bar{t}$	2040 ± 140	2920 ± 180	6880 ± 140
Single-t	370 ± 40	490 ± 60	700 ± 80
SM Diboson	350 ± 90	170 ± 50	51 ± 14
Z+jets	49 ± 6	143 ± 17	15 ± 3
Multijet	–	–	–
Total background	5940 ± 70	10570 ± 100	8190 ± 90
Observed	5885	10619	8178
	Merged, Low Purity		
W+jets	10790 ± 250	10970 ± 260	1420 ± 170
$t\bar{t}$	2650 ± 190	3790 ± 220	8740 ± 240
Single-t	490 ± 60	550 ± 60	820 ± 100
SM Diboson	430 ± 120	200 ± 60	70 ± 20
Z+jets	205 ± 25	215 ± 27	54 ± 9
Multijet	–	–	–
Total background	14567 ± 120	15730 ± 120	11110 ± 100
Observed	14566	15707	11133
	Resolved		
W+jets	61500 ± 1800	165700 ± 2700	8000 ± 900
$t\bar{t}$	23300 ± 1600	31100 ± 2100	78400 ± 1300
Single-t	3800 ± 400	4700 ± 500	5600 ± 700
SM Diboson	2400 ± 700	1500 ± 400	270 ± 100
Z+jets	1750 ± 270	4300 ± 600	280 ± 100
Multijet	3600 ± 700	7600 ± 1700	800 ± 140
Total background	96410 ± 310	214900 ± 500	93280 ± 310
Observed	96459	214838	93257

TABLE 10.3: Expected and observed number of events in the signal and control regions for the WZ channel, evaluated after a background-only fit to the data in the ggF/$q\bar{q}$ category. The uncertainty on estimation of the total background can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources [38].

	WZ Signal region	W+jets Control region	Top Control region
	Merged, High Purity		
W+jets	3680 ± 170	6960 ± 190	560 ± 60
$t\bar{t}$	2280 ± 150	2810 ± 170	6840 ± 140
Single-t	410 ± 50	490 ± 60	750 ± 90
SM Diboson	360 ± 100	160 ± 40	51 ± 14
Z+jets	56 ± 7	148 ± 18	15 ± 3
Multijet	–	–	–
Total background	6780 ± 80	10570 ± 100	8210 ± 90
Observed	6751	10619	8178
	Merged, Low Purity		
W+jets	13400 ± 300	11090 ± 250	1500 ± 170
$t\bar{t}$	3450 ± 230	3680 ± 220	8610 ± 240
Single-t	660 ± 80	560 ± 70	850 ± 100
SM Diboson	500 ± 140	190 ± 50	70 ± 20
Z+jets	240 ± 30	212 ± 26	55 ± 9
Multijet	–	–	–
Total background	18200 ± 140	15730 ± 120	11090 ± 100
Observed	18188	15707	11133
	Resolved		
W+jets	49100 ± 1300	164500 ± 2700	8100 ± 900
$t\bar{t}$	24400 ± 1300	30600 ± 2000	78000 ± 1300
Single-t	3500 ± 400	4700 ± 600	5800 ± 700
SM Diboson	1700 ± 500	1500 ± 400	270 ± 80
Z+jets	1480 ± 260	4400 ± 700	280 ± 60
Multijet	2700 ± 500	9000 ± 1900	900 ± 150
Total background	82720 ± 290	214900 ± 500	93300 ± 300
Observed	82740	214838	93257

TABLE 10.4: Expected and observed number of events in the signal and control regions for the WW channel, evaluated after a background-only fit to the data in the VBF category. The uncertainty on estimation of the total background can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources [38].

	WW Signal region	W+jets Control region	Top Control region
	Merged, High Purity		
W+jets	71 ± 15	183 ± 26	18 ± 4
$t\bar{t}$	84 ± 16	179 ± 22	346 ± 19
Single-t	13 ± 3	24 ± 6	30 ± 5
SM Diboson	10 ± 3	13 ± 4	3.3 ± 1.1
Z+jets	1.6 ± 0.5	4.5 ± 0.9	0.5 ± 0.3
Multijet	–	–	–
Total background	180 ± 12	404 ± 19	398 ± 18
Observed	176	402	398
	Merged, Low Purity		
W+jets	268 ± 31	294 ± 35	55 ± 11
$t\bar{t}$	115 ± 24	225 ± 30	500 ± 27
Single-t	23 ± 5	31 ± 6	47 ± 9
SM Diboson	17 ± 6	16 ± 4	7 ± 3
Z+jets	6.7 ± 2.1	8.7 ± 2.1	2.0 ± 0.7
Multijet	–	–	–
Total background	430 ± 20	575 ± 23	611 ± 23
Observed	436	567	613
	Resolved		
W+jets	1090 ± 110	2520 ± 190	220 ± 50
$t\bar{t}$	710 ± 110	1040 ± 140	2440 ± 90
Single-t	66 ± 16	104 ± 24	120 ± 21
SM Diboson	52 ± 19	66 ± 22	14 ± 6
Z+jets	41 ± 10	94 ± 30	12 ± 4
Multijet	44 ± 19	100 ± 40	54 ± 19
Total background	2010 ± 50	3920 ± 70	2860 ± 50
Observed	2004	3924	2856

TABLE 10.5: Expected and observed number of events in the signal and control regions for the WZ channel, evaluated after a background-only fit to the data in the VBF category. The uncertainty on estimation of the total background can be smaller than the quadratic sum of the individual background contributions due to anti-correlations between the estimates of different background sources [38].

	WZ Signal region	W+jets Control region	Top Control region
	Merged, High Purity		
W+jets	75 ± 17	187 ± 27	18 ± 5
$t\bar{t}$	106 ± 24	170 ± 50	350 ± 40
Single-t	12 ± 6	24 ± 10	30 ± 10
SM Diboson	10 ± 5	11 ± 5	2.7 ± 1.1
Z+jets	1.7 ± 1.5	4.6 ± 2.3	0.5 ± 0.2
Multijet	–	–	–
Total background	205 ± 14	400 ± 20	400 ± 20
Observed	201	402	398
	Merged, Low Purity		
W+jets	320 ± 40	300 ± 40	58 ± 12
$t\bar{t}$	160 ± 50	220 ± 60	500 ± 50
Single-t	26 ± 11	30 ± 9	46 ± 19
SM Diboson	21 ± 10	13 ± 5	6 ± 4
Z+jets	8 ± 6	9 ± 4	1.9 ± 1.2
Multijet	–	–	–
Total background	540 ± 26	580 ± 27	610 ± 30
Observed	550	567	613
	Resolved		
W+jets	770 ± 260	2500 ± 600	190 ± 50
$t\bar{t}$	900 ± 200	1060 ± 260	2470 ± 90
Single-t	80 ± 40	110 ± 60	120 ± 50
SM Diboson	37 ± 23	61 ± 27	12 ± 5
Z+jets	53 ± 15	80 ± 40	11 ± 4
Multijet	30 ± 28	94 ± 40	56 ± 20
Total background	1830 ± 50	3900 ± 80	2860 ± 60
Observed	1829	3924	2856

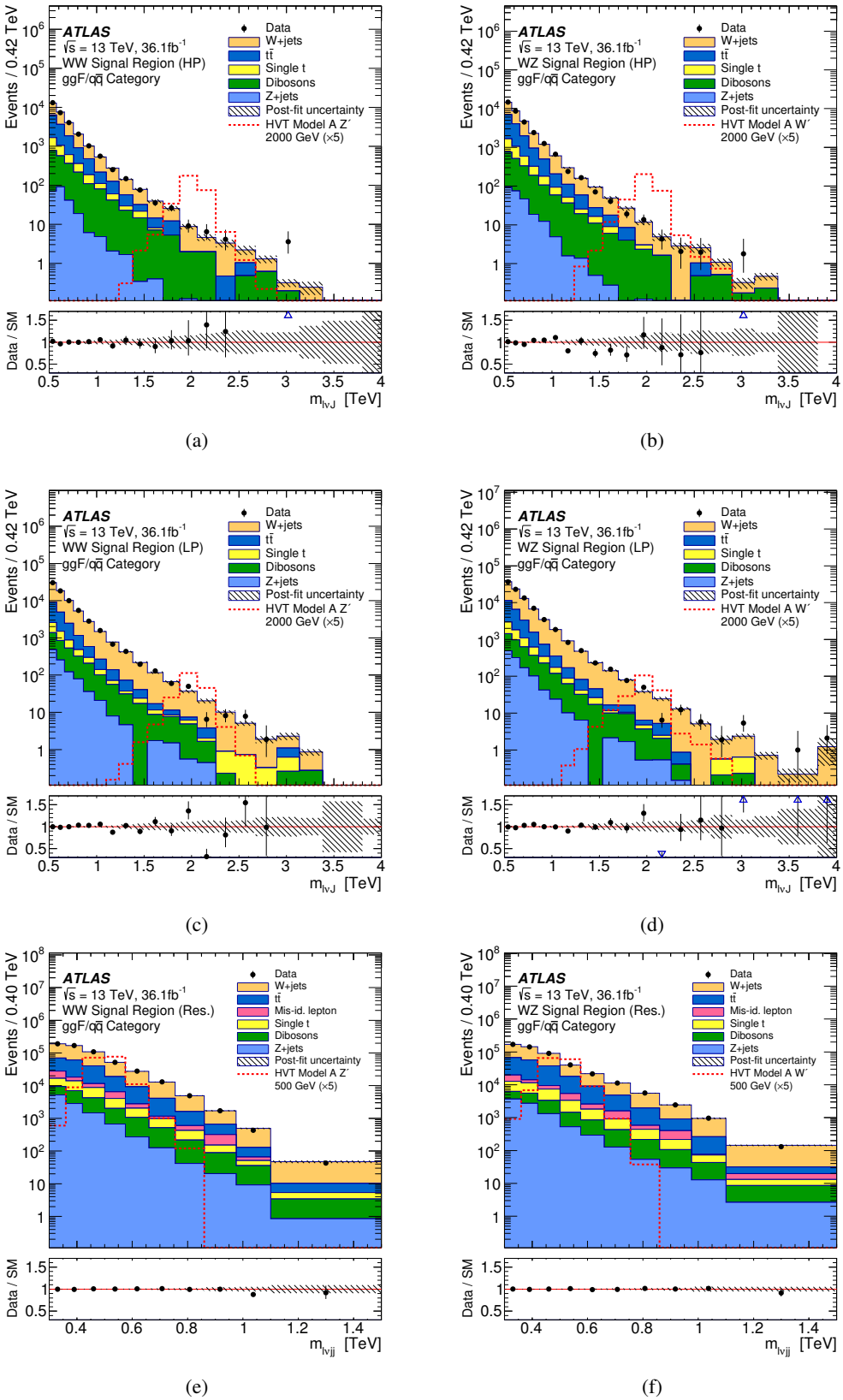


FIGURE 10.2: $m(WV)$ distributions in the Post-fit signal region in the ggF/$q\bar{q}$ category. The merged high-purity (HP) sample of (a) WW and (b) WZ events, the merged low-purity (LP) sample of (c) WW and (d) WZ events and the resolved (Res.) sample of (e) WW and (f) WZ events are presented. The expected background after the profile likelihood fit to the data is shown, and signal predictions are overlaid. The HVT Z'/W' Model A signal at 2000 GeV is presented for the merged analysis, while the 500 GeV signal is shown in the resolved topology. The band denotes the statistical and systematic uncertainty in the background after the fit to the data. The lower panels show the ratio of the observed data to the estimated SM background. In all regions, the number of events is normalised by the width of the second last bin, while the overflow events are included in the last bin [38].

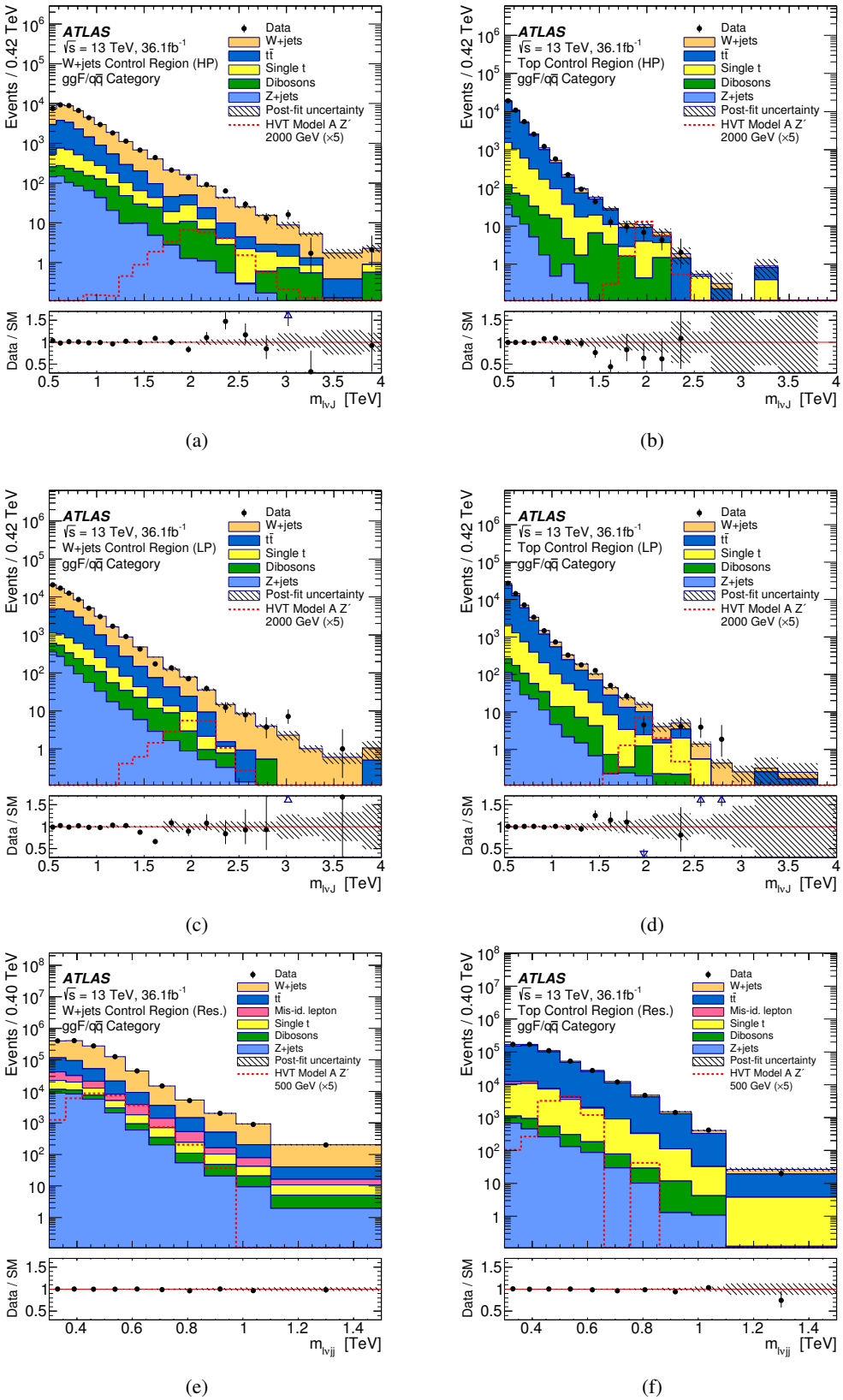


FIGURE 10.3: $m(WV)$ distributions n the post-fit control region in the ggF/qq̄ category. The merged high purity (HP) (a) W+jets and (b) $t\bar{t}$, the merged low purity (LP) (c) W+jets and (d) $t\bar{t}$, and the resolved (Res.) (e) W+jets and (f) $t\bar{t}$ control regions are presented. The background expectation after the profile likelihood fit to the data is shown, and signal expectations are overlaid. The HVT Z'/W' Model A signal at 2000 GeV is presented in the merged channels, while the 500 GeV signal is shown in the resolved topology. The band denotes the statistical and systematic uncertainty on the background after the fit to the data. The lower panels provide the ratio of the observed data to the SM background estimation. In all regions, the number of events is normalised by the width of the second last bin, while the overflow events are included in the last bin [38].

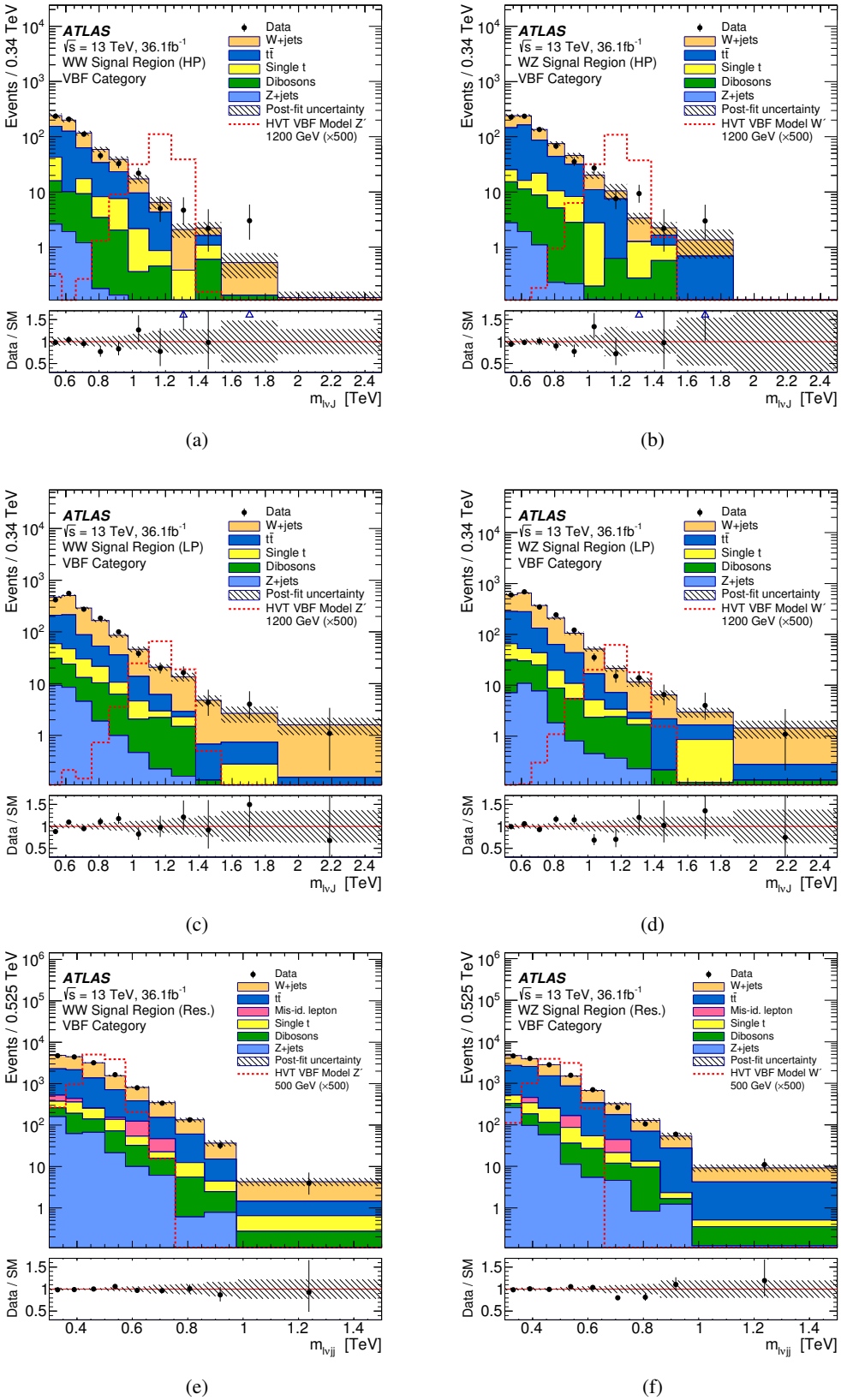
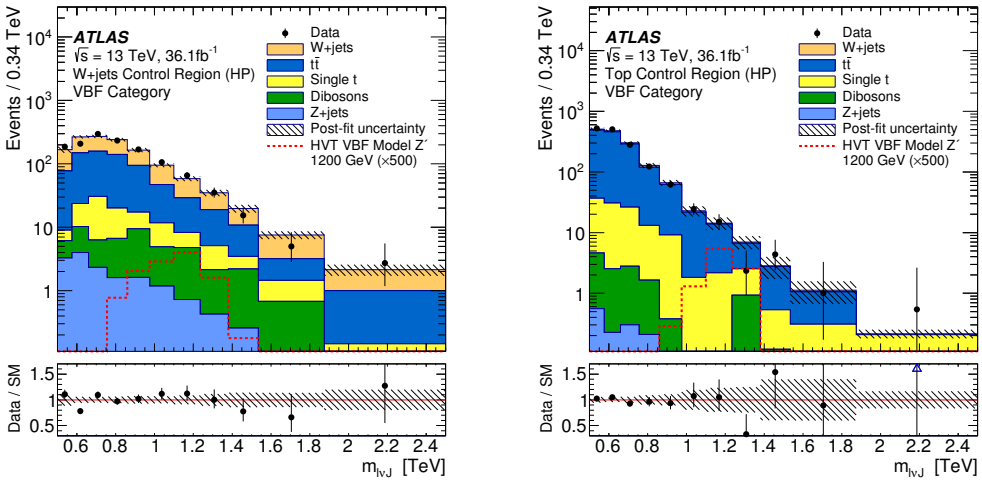
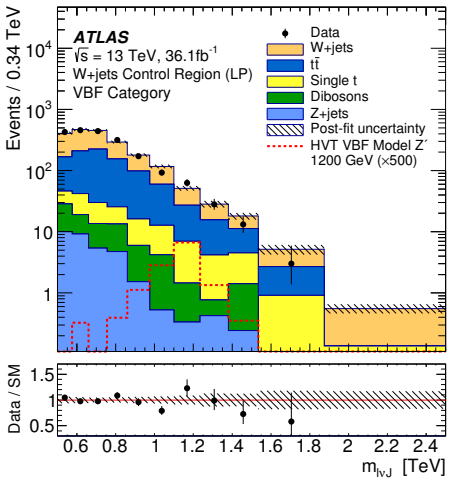


FIGURE 10.4: $m(WV)$ distributions in the post-fit signal region in the VBF category. The merged high-purity (HP) sample of (a) WW and (b) WZ events, the merged low-purity (LP) sample of (c) WW and (d) WZ events and the resolved (Res.) sample of (e) WW and (f) WZ events are presented. The expected background after the profile likelihood fit to the data is shown, and signal predictions are overlaid, normalised to the cross sections indicated in the legends. The VBF HVT Z'/W' signal at 1200 GeV is presented for the merged analysis, while the 500 GeV signal is shown in the resolved topology. The band denotes the statistical and systematic uncertainty in the background after the fit to the data. The lower panels show the ratio of the observed data to the estimated SM background. In all regions, the number of events is normalised by the width of the second last bin, while the overflow events are included in the last bin [38].

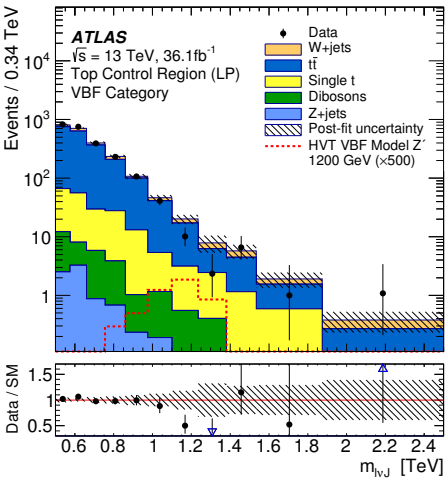


(a)

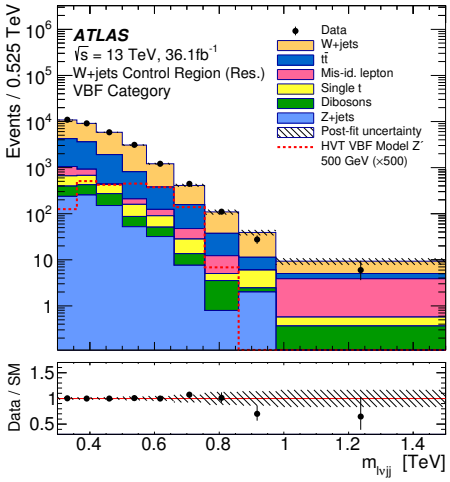
(b)



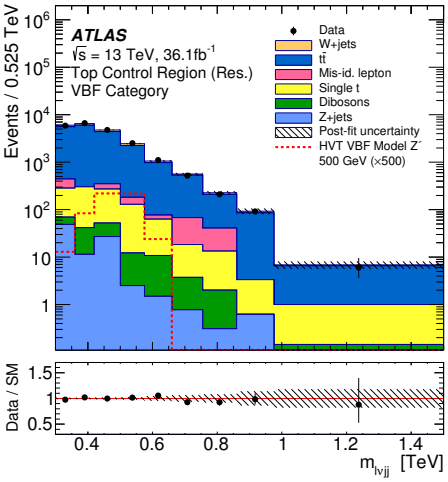
(c)



(d)



(e)



(f)

FIGURE 10.5: $m(WV)$ distributions in the post-fit control region in the VBF category. The merged high purity (HP) (a) W+jets and (b) $t\bar{t}$, the merged low purity (LP) (c) W+jets and (d) $t\bar{t}$, and the resolved (Res.) (e) W+jets and (f) $t\bar{t}$ control regions are presented. The background expectation after the profile likelihood fit to the data is shown, and signal expectations are overlaid, normalised to the cross sections indicated in the legends. The VBF HVT Z'/W' signal at 1000 GeV is presented in the merged channels, while the 500 GeV signal is shown in the resolved topology. The band denotes the statistical and systematic uncertainty on the background after the fit to the data. The lower panels provide the ratio of the observed data to the SM background estimation. In all regions, the number of events is normalised by the width of the second last bin, while the overflow events are included in the last bin [38].

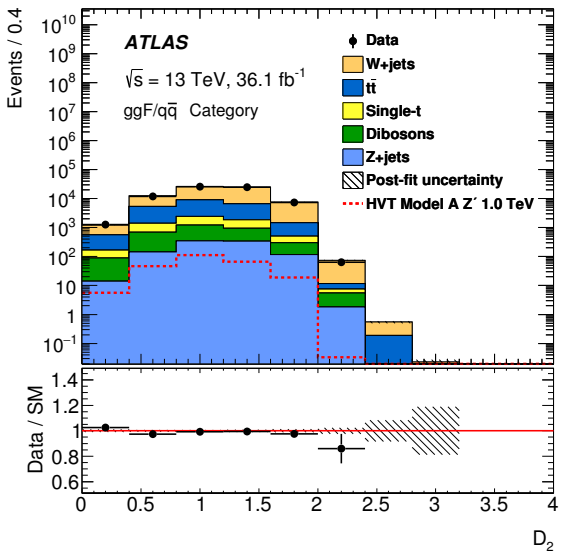


FIGURE 10.6: Distribution of the D_2 variable, which is used to distinguish the high- and low-purity regions. The requirement on the D_2 value depends on the transverse momentum of the jet. Post-fit normalisation factors are used to normalise the background contributions and the HVT WW signal sample with $m(Z') = 1.0$ TeV is overlaid in the ggF/qq̄ category. The band denotes the statistical and systematic uncertainty on the background after the fit to the data, while the lower panel shows the ratio of the observed data to the SM background estimation [38].

Table 10.6 summarises exclusion limits on the mass extracted from the ggF/qq̄ category for the various signal hypotheses. In Figures 10.9 and 10.11, the upper limits for intermediate mass points where simulated signal samples are not available are evaluated as a linear assumption of the neighbouring mass points (interpolation). The interpolation to the shapes of the signal distributions using is performed the moment-morphing method [155], and confirmed that the linear assumption works fine in this analysis. The details are described in the Appendix B.

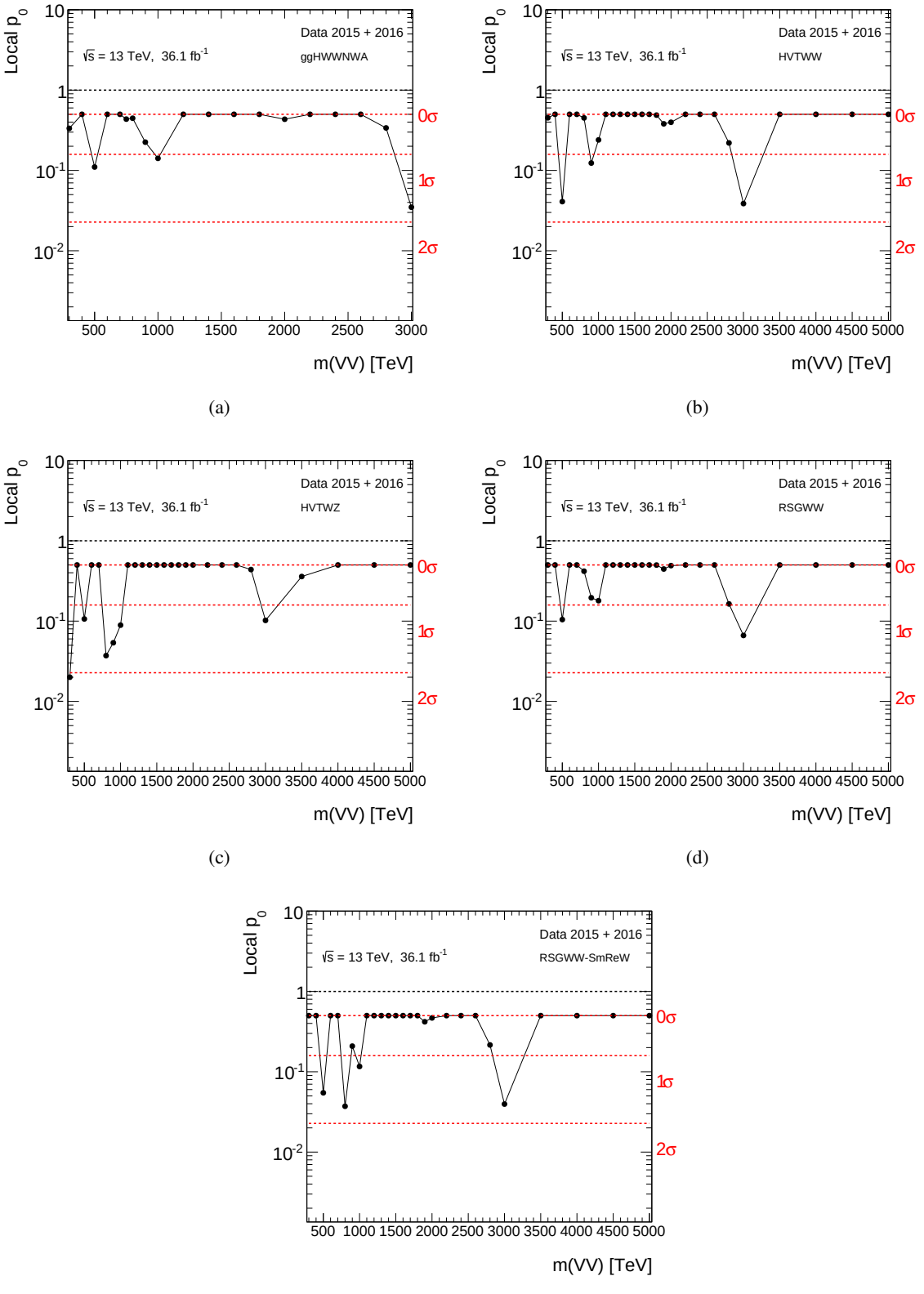
10.5 Impacts of the systematic uncertainties

Lists of leading sources of uncertainties in the best-fit for the μ value ($\Delta\mu/\mu$) is presented in Table 10.7. Examples from the high and low mass regions, where the highest sensitivity is achieved in the merged and the resolved topologies respectively, in the ggF/qq̄ and VBF categories. The statistical uncertainty in data is dominant at the mass range of $m_{WV} > 700$ GeV.

TABLE 10.6: Observed and expected excluded masses at the 95% confidence level for various signal hypotheses in the ggF/qq̄ category [38].

WW Selection			
Excluded	HVT		RS G_{KK}
Masses	Model A	Model B	$k/\bar{M}_{Pl} = 1.0$
Observed	<2750 GeV	<3000 GeV	<1750 GeV
Expected	<2850 GeV	<3150 GeV	<1750 GeV

WZ Selection		
Excluded	HVT	
Masses	Model A	Model B
Observed	<2800 GeV	<3000 GeV
Expected	<2900 GeV	<3200 GeV



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

FIGURE 10.7: Local p_0 values as a function of the invariant mass of two vector bosons in the ggF/$q\bar{q}$ category, for the signals; (a) scalar $H \rightarrow WW$, (b) HVT Z', (c) HVT W', (d) G_{KK} with $k/\bar{M}_{Pl} = 1.0$, and (e) G_{KK} with $k/\bar{M}_{Pl} = 0.5$.

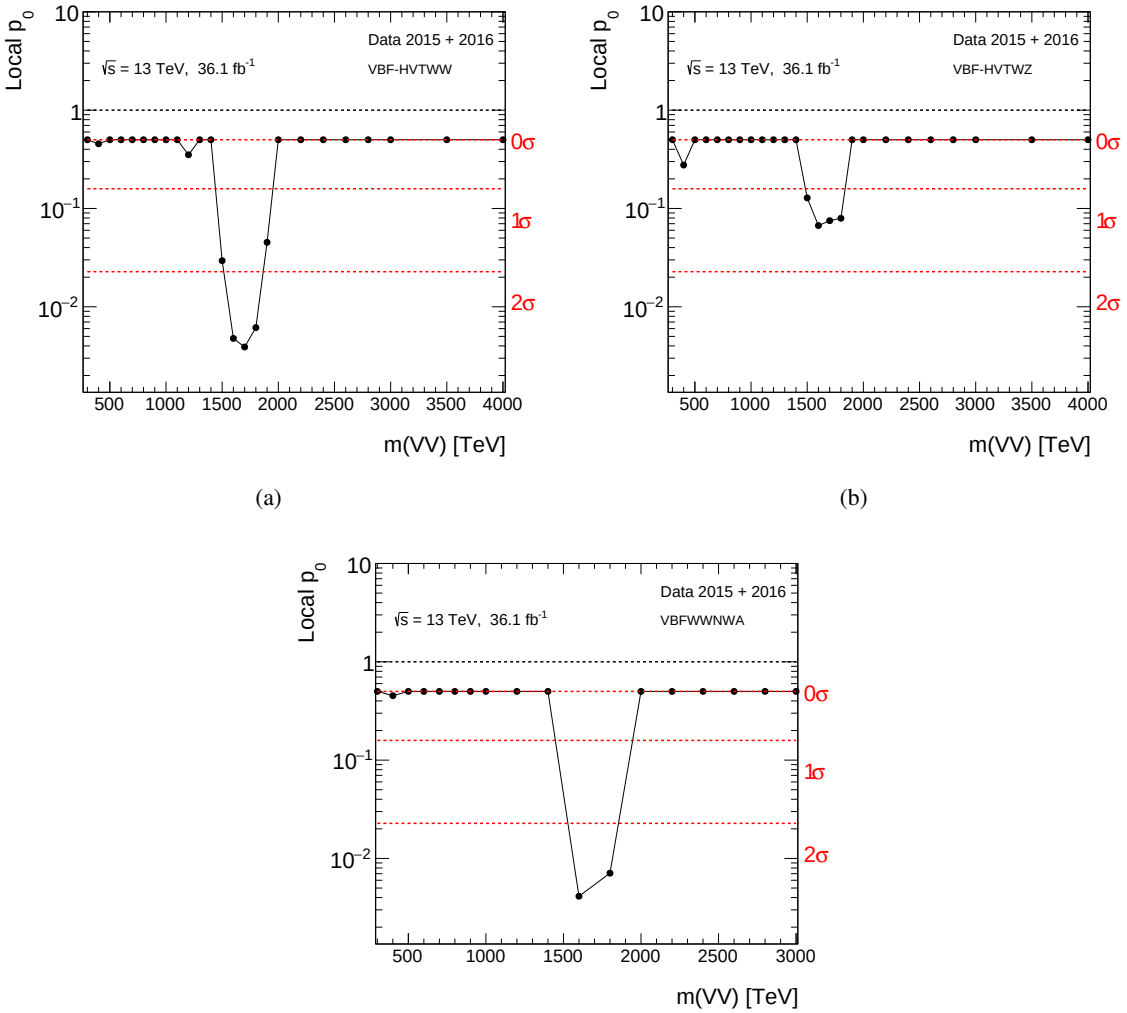


FIGURE 10.8: Local p_0 values as a function of the invariant mass of the two vector bosons in the VBF category, for the signals; (a) HVT Z', (b) HVT W', and (c) heavy scalar signals, H, produced via VBF.

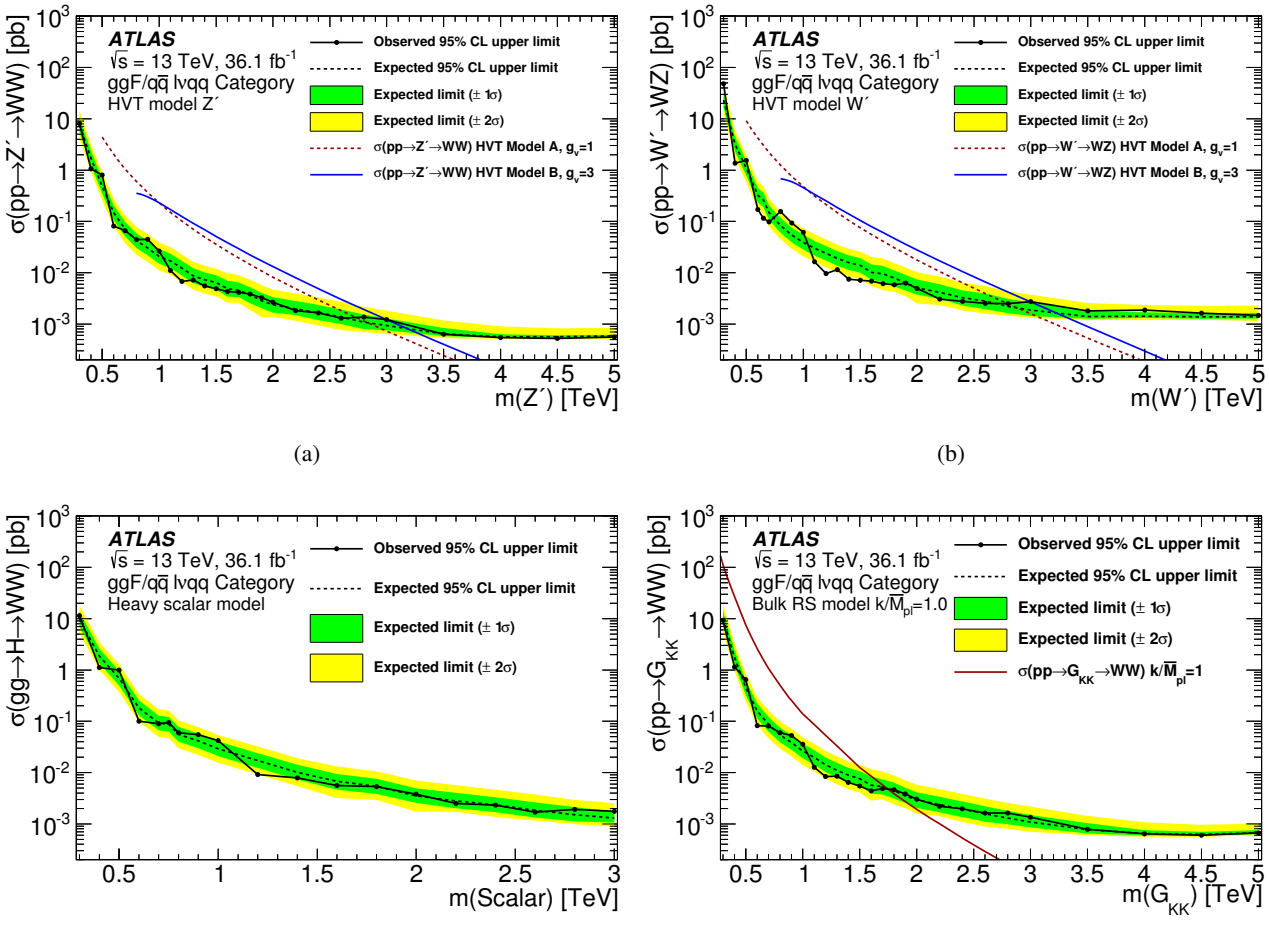


FIGURE 10.9: The observed and expected upper limits for the cross-section including the branching fractions of the decay to the WV final state at the 95% confidence level in the ggF/$q\bar{q}$ category are presented [38]. The upper limits are shown as a function of the resonance mass. Interpretations for (a) HVT WW, (a) HVT WZ, (c) scalar $H \to WW$, and (d) G_{KK} are presented. The red and blue curves show the predicted signal cross-sections as a function of the invariant mass. The scalar $H \to WW$ signal is used for a feasibility study of searches for extended Higgs sectors.

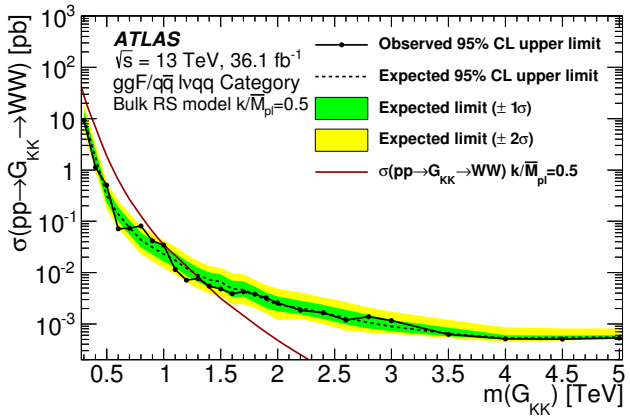


FIGURE 10.10: The observed and expected upper limits for the cross-section including the branching fractions of the decay to the WV final state at the 95% confidence level via gluon-gluon fusion are presented [38]. The upper limits are shown as a function of the resonance mass. To compare the results with the CMS results, the interpretation for $pp \to G_{\mathrm{KK}} \to WW$ production with $k/\bar{M}_{\mathrm{Pl}} = 0.5$ is presented. Masses below 1.35 TeV are excluded at 95% CL, except for the regions between 0.98 TeV to 1.02.

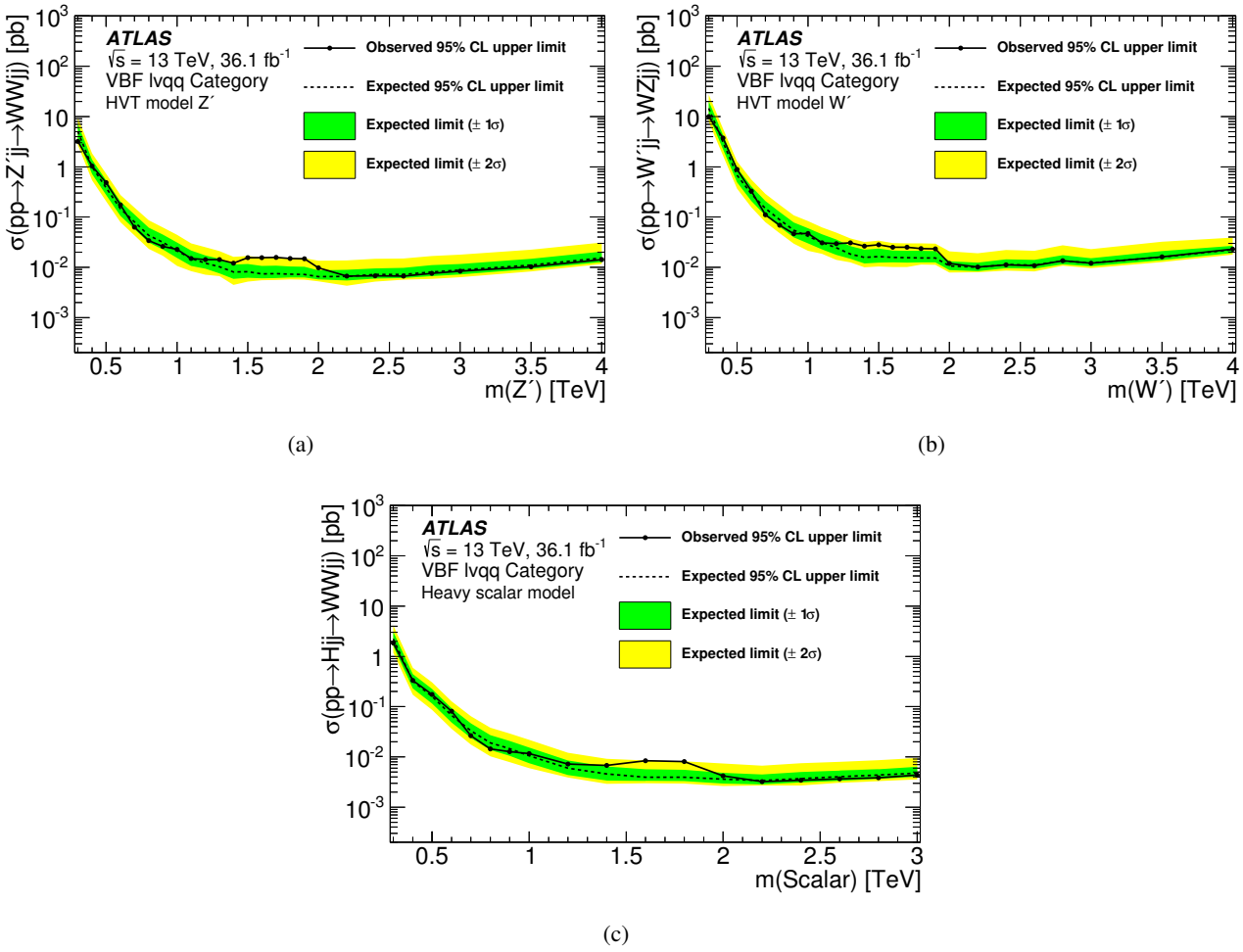


FIGURE 10.11: The observed and expected upper limits for the cross-section including the branching fractions of the decay to the WV final state at the 95% confidence level in the VBF category are presented [38]. The upper limits are shown as a function of the resonance mass. Interpretations for (a) HVT Z', (b) HVT W' and (c) heavy scalar signals, H, produced via VBF are shown. There is a slight excess around 1.5 TeV. The mass region greater than 1.5 TeV is covered by two bins in $m(WV)$ as described in Section 9.1, while the observed limit dots in this figure represent the resonance mass values considered in the interpretation.

At the high-mass region, the dominant systematic uncertainties come from the W+jets and the $t\bar{t}$ modelling as well as the large-R jets mass resolution uncertainty, with 10–20 % on the best fit signal μ value; at the low-mass region, the dominant systematic uncertainties come from the uncertainties related to the background modelling as well as the uncertainty of the large-R jet energy calibration. In Section 10.2-10.5, the latest results of the large-R JES calibration, described in Chapter 5, are not used, thus the uncertainties related to the large-R jet are large; 17 % uncertainty at the signal mass of 500 GeVin the ggF/$q\bar{q}$ category. The uncertainties from the small-R jet energy resolution are large as well; 9 % uncertainty at the signal mass of 500 GeVin the VBF category.

10.6 Future prospects; expected improvements by applying the latest large-R jet energy scale calibrations

The latest large-R jet energy calibrations described in Section 5.2 and Reference [48], are not used in the diboson resonance search presented in Section 10.2-10.5. In this section, expected improvements by applying the new large-R jet energy scale calibrations and expected improvements with the full data taken in Run-2 are described; we would apply the latest calibrations to the analysis with the full Run-2 dataset, corresponding to a total integrated luminosity of about 120 fb^{-1}[2]. The the expected upper limits presented in this section are, hence, obtained in the merged analysis and normalised to the expected full Run-2 statistics; then we compare the calculated limits with the limits calculated with the latest large-R jet energy scale calibrations.

The larger datasets will allow us to extend the excluded masses, however, as we see in Table 10.7, the systematic uncertainties are dominant in the low mass region (mass$^{\mathrm{BSMparticle}} \lesssim 700$ GeV) whereas the statistical uncertainty in data is dominant at the high mass region; it is important to increase the sensitivity to diboson resonances by using the larger datasets, however we also need to pay attention to the techniques by which we can decrease the systematic uncertainties.

In Section 10.2-10.5, the value $k/\bar{M}_{\mathrm{Pl}} = 1.0$ is used to interpret the G_{KK} resonance, although the $k/\bar{M}_{\mathrm{Pl}}$ of order 0.1 or less is preferred theoretically [13]. Note that both the production cross section and decay width of the KK graviton scale as the square of $k/\bar{M}_{\mathrm{Pl}}$. By decreasing the systematic uncertainties, we can set the upper limits to the G_{KK} resonance with smaller $k/\bar{M}_{\mathrm{Pl}}$ values. Furthermore, as described in Section 1.2.1, our final goal is to test the electroweak symmetry breaking in the vector boson scattering processes. Improving the identification of boosted W and Z bosons by applying the latest large-R jet energy calibrations, we can start vector boson scattering searches in the semi-leptonic final state.

Expected upper limits with the latest large-R jet energy scale calibrations

The expected upper limit for the cross section of HVT $Z' \rightarrow WW$, G_{KK}, and the heavy scalar signal in the ggF/$q\bar{q}$ category are shown in Figure 10.12, 10.13, and 10.14. In the bottom panels, the ratio between the upper limits calculated with the old large-R jet energy scale systematic uncertainties and with the latest large-R jet energy scale calibrations are shown as a function of the invariant mass of $\ell\nu J$. It is confirmed that the expected upper limits for the new particles with the mass of 500 GeV would improve by 8-10 % compared to the currently used systematic uncertainties.

Impacts on the BSM searches

The current sensitivity to the HVT $Z' \rightarrow WW$ search is higher in the final states of $\ell\nu\ell\nu/\ell\ell qq$ in the low mass region than the sensitivity of the search carried out in the final state of $\ell\nu qq$ by 8-10 %. By improving the limits in the $\ell\nu qq$ final state by 10% in the low mass region, the upper limits of the $\ell\nu qq$ would have the highest sensitivity to the HVT $Z' \rightarrow WW$ search, and the combined results of the $Z' \rightarrow WW$ searches would improve significantly; we can also improve the combined results of the $VV \rightarrow WW/WZ/ZZ$ searches significantly by applying the latest large-R jet energy scale calibrations[3].

Even using the latest large-R jet energy calibration and the full Run-2 dataset, we can not set upper limits to the G_{KK} resonance with a theoretically favoured $k/\bar{M}_{\mathrm{Pl}}$ value (e.g. $k/\bar{M}\mathrm{Pl} = 0.1$, the production cross section of

[2]LHC Run-2 proton-proton collision program finished on 29th October 2018, the total integrated luminosity certified good quality for physics analyses is 140 fb^{-1} [59]

[3]The latest combination of searches for heavy diboson resonances are described in Reference [158]

TABLE 10.7: Dominant relative uncertainties in the signal-strength parameter (μ). The uncertainties are calculated with the following hypotheses; HVT signal production with $m(Z') = 1200$ GeV and $m(W') = 500$ GeV in the VBF category, and $m(W') = 2000$ GeV and $m(Z') = 500$ GeV in the ggF/$q\bar{q}$ category, assuming that the production cross sections equal to the expected 95% CL upper limits of 0.012 pb, 0.7 pb, 0.005 pb, and 0.5 pb, respectively. The impact from the many other sources of systematic uncertainty remains significant. The effect of the statistical uncertainty on the signal and background samples is also shown. The kinematic uncertainties on the large-R jets arise from jet reconstruction uncertainties that can be dominant in the low $m(WV)$ region because of the merged analysis priority in the event categorisation. The scale uncertainty of the $t\bar{t}$ background includes the uncertainties of the factorisation and renormalization scales of the nominal generator. The scale uncertainty of the W+jets background includes the uncertainties in the renormalization and factorisation scales, the CKKW matching scales[1] [156, 157], and the resummation scale. The cross section uncertainties for the W+jets and $t\bar{t}$ backgrounds are constrained by the corresponding control data [38].

VBF Category			
$m(Z') = 1200$ GeV		$m(W') = 500$ GeV	
Source	$\Delta\mu/\mu$ [%]	Source	$\Delta\mu/\mu$ [%]
MC statistical uncertainty	15	MC statistical uncertainty	16
Large-R jets mass resolution	5	W+jets: cross section	10
W+jets: PDF choice	5	Multijet $E_{\mathrm{T}}^{\mathrm{miss}}$ modelling	10
$t\bar{t}$: alternative generator	5	Small-R jets energy resolution	9
W+jets: cross section	5	SM diboson cross section	8
$t\bar{t}$: scales	4	$t\bar{t}$: cross section	7
Total systematic uncertainty	24	Total systematic uncertainty	40
Statistical uncertainty	52	Statistical uncertainty	30

ggF/$q\bar{q}$ Category			
$m(W') = 2000$ GeV		$m(Z') = 500$ GeV	
Source	$\Delta\mu/\mu$ [%]	Source	$\Delta\mu/\mu$ [%]
MC statistical uncertainty	12	Large-R jets kinematics	17
W+jets: generator choice	8	MC statistical uncertainty	12
W+jets: scale	5	$t\bar{t}$: scale	11
SM diboson normalisation	4	SM diboson cross section	10
Large-R jets mass resolution	4	W+jets: alternative generator	10
Large-R jets D_2 resolution	4	W+jets: scale	9
Total systematic uncertainty	20	Total systematic uncertainty	42
Statistical uncertainty	50	Statistical uncertainty	18

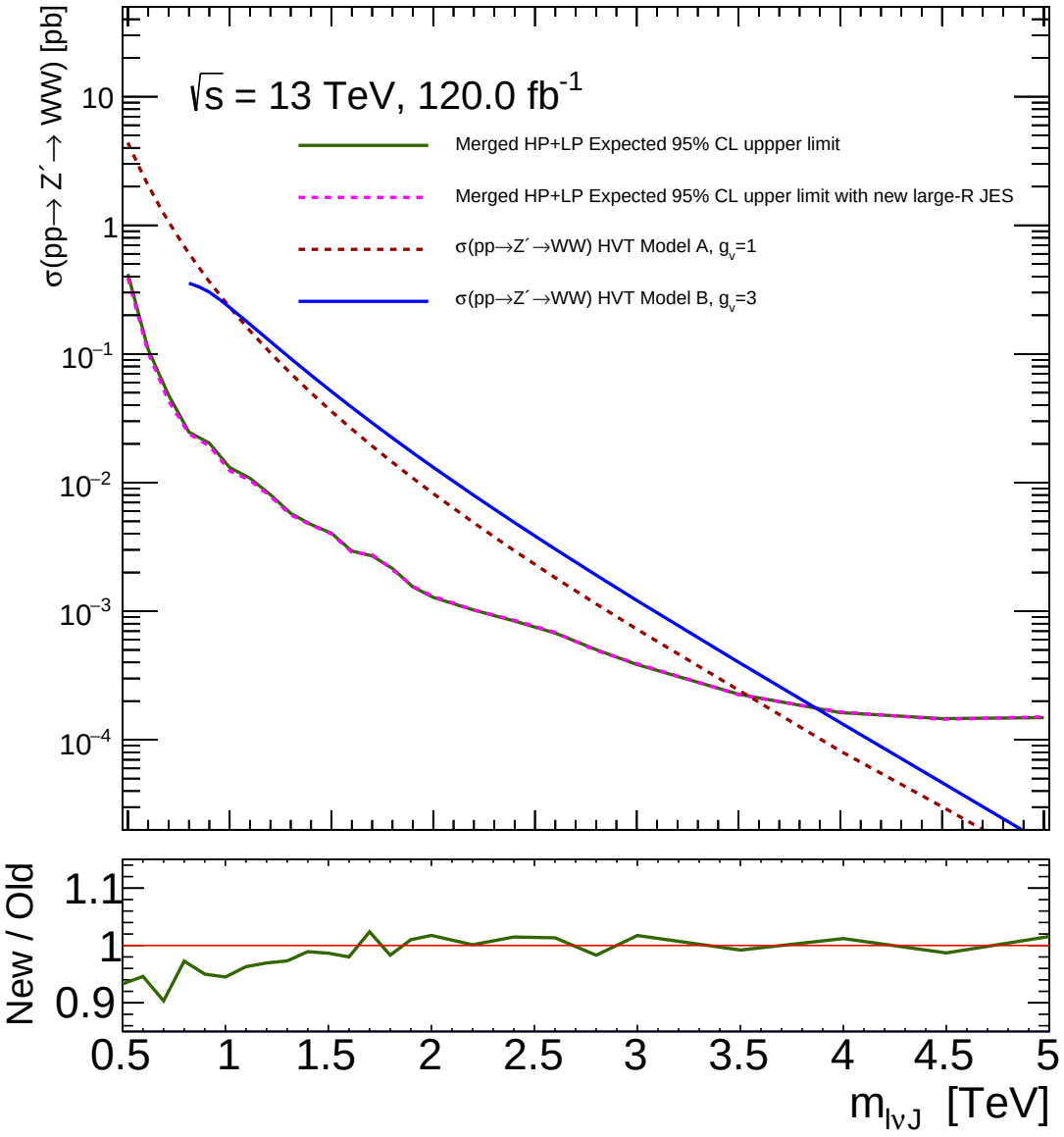


FIGURE 10.12: Expected upper limits for the HVT $Z' \to WW$ signal in the ggF/$q\bar{q}$ category carried out by the merged analysis with the total integrated luminosity of 120 fb^{-1}. The green solid line shows the expected upper limits at the 95 % confidence-level obtained with the same systematic uncertainties used in the results shown in Section 10.2-10.5. The magenta dashed line shows the expected upper limits with the latest large-R jet energy scale calibrations, described in Chapter 5. The bottom panels shows the ratio between the upper limits calculated with the old large-R jet energy scale systematic uncertainties and with the latest large-R jet energy scale calibrations, as a function of the invariant mass of $\ell\nu J$. At the low mass region, it is confirmed that the expected upper limits is improved by 8-10 % by applying the latest large-R jet energy scale calibrations.

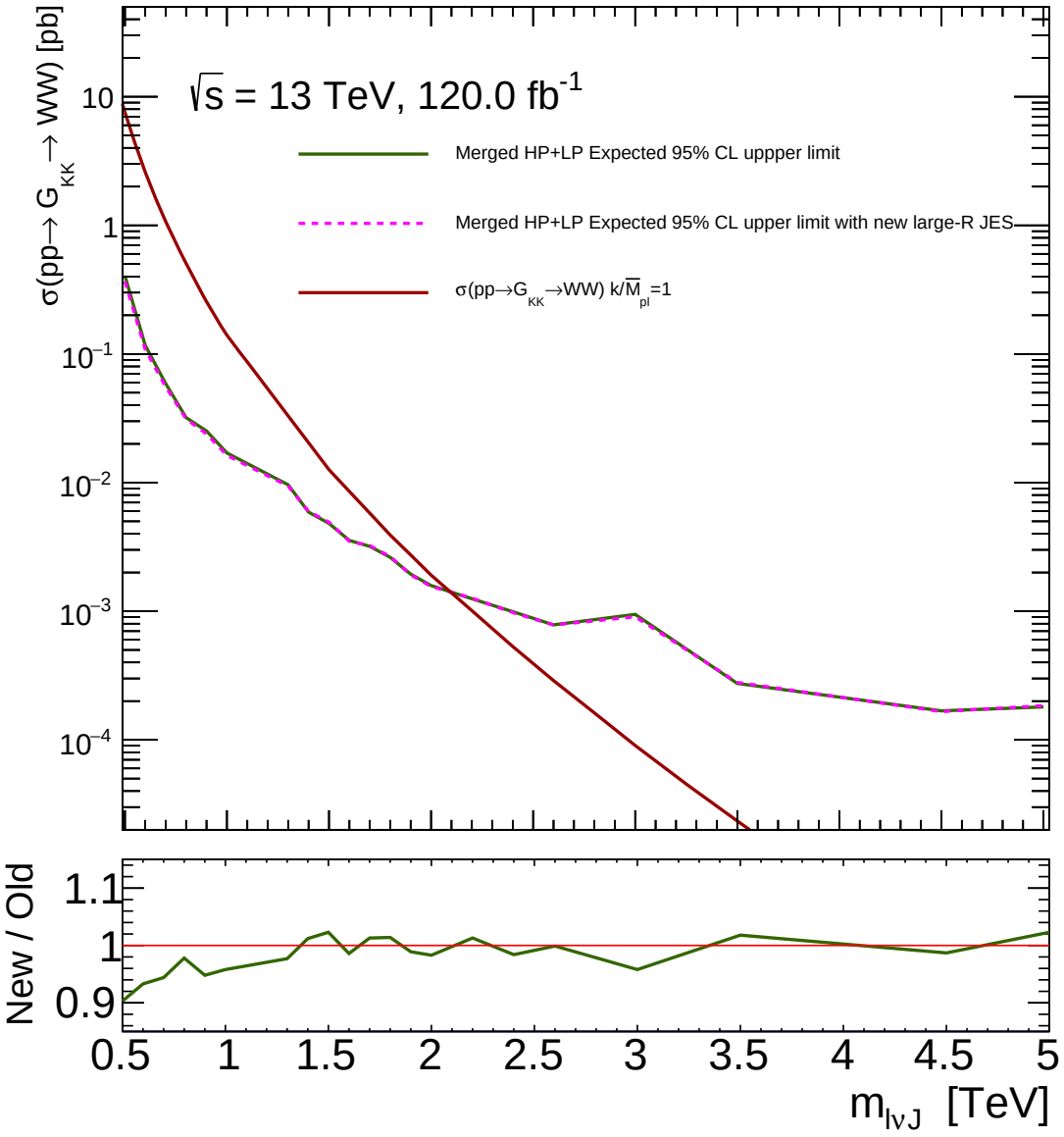


FIGURE 10.13: Expected upper limits for the G_{KK} signal in the ggF/$q\bar{q}$ category carried out by the merged analysis with the total integrated luminosity of 120 fb^{-1}. The green solid line shows the expected upper limits at the 95 % confidence-level obtained with the same systematic uncertainties used in the results shown in Section 10.2-10.5. The magenta dashed line shows the expected upper limits with the latest large-R jet energy scale calibrations, described in Chapter 5. The bottom panels shows the ratio between the upper limits calculated with the old large-R jet energy scale systematic uncertainties and the latest large-R jet energy scale calibrations, as a function of the invariant mass of $\ell\nu J$. At the low mass region, it is confirmed that the expected upper limits is improved by 8-10 % by applying the latest large-R jet energy scale calibrations.

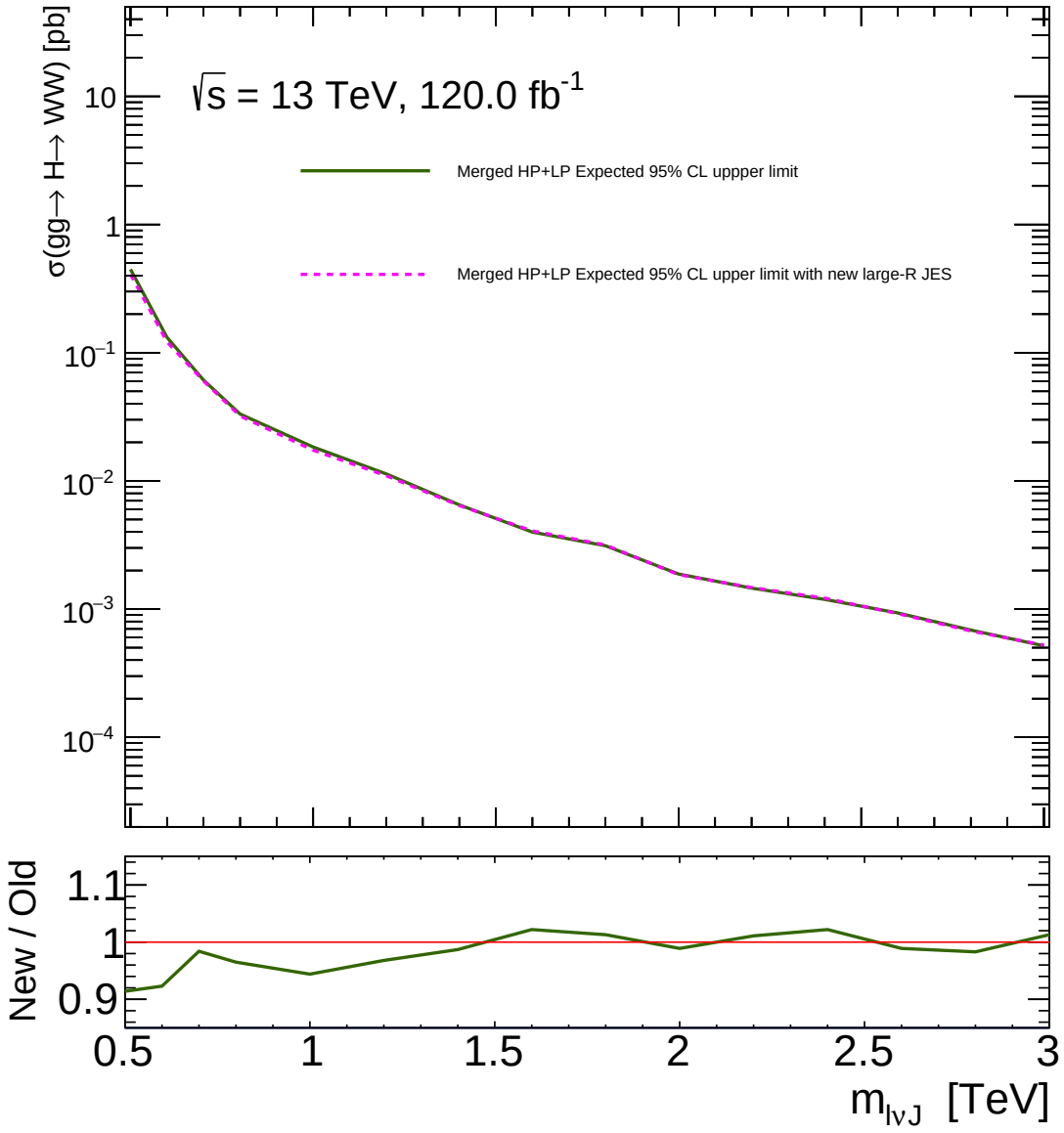


FIGURE 10.14: Expected upper limits for the heavy scalar signal in the ggF/$q\bar{q}$ category carried out by the merged analysis with the total integrated luminosity of 120 fb^{-1}. The green solid line shows the expected upper limits at the 95 % confidence-level obtained with the same systematic uncertainties used in the results shown in Section 10.2-10.5. The magenta dashed line shows the expected upper limits with the latest large-R jet energy scale calibrations, described in Chapter 5. The bottom panels shows the ratio between the upper limits calculated with the old large-R jet energy scale systematic uncertainties and the latest large-R jet energy scale calibrations, as a function of the invariant mass of $\ell\nu J$. At the low mass region, it is confirmed that the expected upper limits is improved by 8-10 % by applying the latest large-R jet energy scale calibrations.

the G_{KK} resonances scales 0.01 with this value.). To search the G_{KK} resonances effectively, much larger dataset is needed, furthermore we need to decrease the systematic uncertainties; the update of the large-R jet energy calibration is a first step of the reduction procedure of the systematic uncertainties.

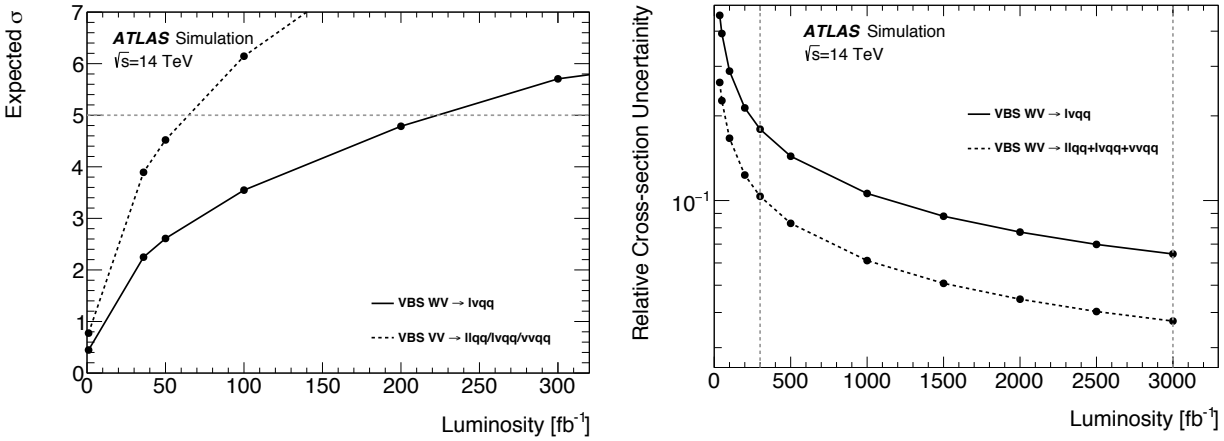
In Section 10.2-10.5, we presented the upper limits on the neutral heavy higgs as a feasibility test for extended Higgs sectors. The neutral heavy higgs with mass of 500 GeV in the low $\tan\beta$ regime (e.g. $\tan\beta = 2.5$) would decay into the WW final state with the branching ratio of $\sim 2\%$ [159], thus we will be able to increase the feasibility of carrying out the searches of the heavy higgs decay $H \to WW$ in the ATLAS experiment by improving the upper limits in the low mass region.

We carried out various BSM searches using the early part of the LHC Run-2 dataset, however, we have not found any significant excesses yet. Now we are proceeding to the next stage of BSM searches; the precise measurements of the vector boson scattering as a final test of the electroweak symmetry breaking. In Chapter 5.2, we review a set of data-driven jet energy calibration methods. By calibrating large-R jet momentum precisely by these methods, we can exploit the semi-leptonic final state in the vector boson scattering search.

The current VBS search in the $\ell\nu qq$ final state is presented in Reference [160]. In this analysis, the large-R jet energy calibration has an uncertainty of approximately 20% on the total background yields. As an extrapolation from our analysis, we expect the uncertainty on the total background yields will be reduced by ~4% in the VBS search applying the large-R jet energy calibration presented in Chapter 5.2.

Further VBS search is currently in progress using the data corresponding to an integrated luminosity of 36.1 fb^{-1} collected at centre of mass energy of $\sqrt{s} = 13$ TeV by the ATLAS detector at the LHC during the 2015 and 2016 runs. In this analysis, the largest source of the systematic uncertainties arises from the background modelling; the second largest source is the uncertainties related to the calibration of large-R jet energy. As a first approximation, we estimate the expected improvements in this analysis by applying the jet calibration from the improvements in the resonance search produced with the VBF process. The jet calibration will improve the uncertainty in the VBS cross-section by approximately 1 %. After this analysis, the VBS search using the full Run-2 dataset will be carried out. For this analysis, much larger MC samples will be produced compared to our analysis, thus the BG modelling will be significantly improved. Therefore, the calibration of the large-R jet energy will play a key role in the VBS search using the full Run-2 dataset.

The LHC Run-2 program came to an end on 2nd December 2018. During the Run-2, we have accumulated the data corresponding to an integrated luminosity of ~150 fb^{-1}; in order to further increase its discovery potential, LHC will be upgraded in two steps. The next running period (Run-3) will start from 2021, and will accumulate 150 fb^{-1}. (i.e. total integrated luminosity in Run-2 + Run-3 is approximately 300 fb^{-1}) After Run-3, LHC will be upgraded to the High Luminosity LHC (HL-LHC) to achieve instantaneous luminosities a factor of five larger than the LHC nominal value. During the HL-LHC period, we will accumulate ~3000 fb^{-1}. Expected results in the VBS search using the HL-LHC dataset are summarised in Figure 10.15. At the integrated luminosity of 3000 fb^{-1}, the relative uncertainty on the VBS cross-section will be 3-4 %. Then, we can test whether the observed value agrees with the SM prediction; in other words, we will be able to answer a critical question "Are there more Higgs bosons?" by the vector boson scattering analysis.



(a) (b)

FIGURE 10.15: (a) Expected signal significance as a function of integrated luminosity up to 300 fb^{-1}. The solid black curve is the significance from the $\ell\nu qq$ channel, while the black dashed curve shows the expected significance from all semi-leptonic channels assuming equal sensitivity. The grey dashed curve highlights the 5σ value. (b) The expected cross-section uncertainty as a function of integrated luminosity up to 3000 fb^{-1}. The solid black curve is the uncertainty from the $\ell\nu qq$ channel, while the dashed curve shows the expected uncertainty from all semi-leptonic channels assuming equal sensitivity. The grey dashed curve highlights the values at 300 fb^{-1} and 3000 fb^{-1} [161].

Chapter 11

Conclusions

After the discovery of the Higgs boson at LHC, we are focusing on solving some problems which can not be explained by the SM; one of the most critical question is to test whether the Higgs boson is the only responsible for unitarity restoration for the vector boson scattering (VBS). The recent results from the ATLAS and CMS collaborations exploit only the fully-leptonic final state, and are still dominated by the statistical uncertainty in data even using the LHC full data. Towards the semi-leptonic final state in future VBS searches, we established the methodology to search for resonant vector bosons. The key technique is large-R jet energy calibration which improves reconstruction efficiency of the hadronically decaying W and Z bosons which are highly-boosted. The search for vector boson resonance is also critical to study many BSM theories; such as composite Higgs models, warped extra dimensions, models with an extended Higgs sector, and grand unified theories, which predict new particles decaying into a pair of vector bosons.

This thesis presents a search for resonant WW and WZ production in the semi-leptonic ($\ell\nu qq$) final state, using the data corresponding to an integrated luminosity of 36.1 fb^{-1} collected at centre of mass energy of $\sqrt{s} = 13$ TeV by the ATLAS detector at the LHC during the 2015 and 2016 runs. The analysis is carried out in different kinematic topologies according to the reconstruction of the hadronically decaying W/Z boson and is optimised for various production mechanisms of the resonance. We develop a set of selections for each topology to maximise the sensitivity to signal models. After the event selections, the major contribution from the SM background arise from top quark pair production and W boson production associated with hadron jets. Contribution of these SM backgrounds in the signal regions are estimated by dedicated analysis for a set of control regions.

We obtain the result that the data are consistent with the SM background hypothesis. Limits on the production cross-section times branching ratio are obtained as a function of the resonance mass arising from three different models; scalar bosons, Heavy Vector Triplet, and RS gravitons. We consider two different production modes; the vector boson fusion and the gluon-gluon or quark-quark fusion. Masses below 2750 (3000) GeV for HVT WW Model-A (B), 2800 (3000) GeV for HVT WZ Model-A (B), and 1750 GeV for RS G_{kk} signals with $k/\bar{M}_{\mathrm{Pl}} = 1.0$ produced via gluon-gluon fusion are excluded at 95% CL.

The LHC Run-2 program came to an end on 2nd December 2018, we will present new results of vector boson resonance searches using the full LHC data. The larger datasets will allow us to extend the excluded masses, furthermore, by applying the latest large-R jet energy calibrations, described in Chapter 5, we will improve the sensitivity at the low-mass region ($m < 700$ GeV); this improvement is critical to searches for some BSM particles, such as heavy neutral Higgs bosons and the KK gravitons. In the final part of this thesis, the expected improvements by applying the new large-R jet energy scale calibrations are estimated. We confirm that the expected upper limits for the new particles with the mass of 500 GeV will improve by 8-10 % compared to the currently used systematic uncertainties. Furthermore we will be able to use the semi-leptonic final state in the searches for VBS by updating the large-R jet energy scale calibrations. The uncertainties related to the calibration of the large-R jet energy will become much more important in the analysis using the full Run-2 dataset.

Appendix A

Mistimed events

Both the 2015 dataset and the 2016 dataset have a small number of mistimed events, which are recovered from a dedicated stream, and have corrected calorimeter energy although tracking and muon information in these events are incorrect. The mistimed events need to be confirmed to have little effect on the analysis.

These events have incorrect tracking and muon information, thus the normal event selections, described in Section.6, can not be applied. (For example, electrons appear as photons in the mistimed events) The impact of the mistimed events on the analysis have been evaluated without using tracking and muon information, imitating the event selections as close as possible.

- At least one large-R jet: the number of large-R jets >= 1
- Large-R jet selection: $p_{\mathrm{T}}^{\mathrm{Jcalo}} > 200$ GeV, $|\eta^{\mathrm{J}}| < 2.0$
- Boson-tag: (LP and HP) 50 GeV $< m_{\mathrm{calo}}^{\mathrm{J}} <$ 150 GeV, (HP) the recommended D_2-cut
- Photon selection: photon-ID “LooseLH”, isolation-cut “FixedCutTightCaloOnly”, $p_{\mathrm{T}}^{\gamma} > 27$ GeV
- Event topology: $|\Delta\phi(\mathrm{J}, \gamma)| > 2.0$

are required, where $\Delta\phi(a, b)$ is the distance in the coordinate between objects a and b.

Table A.1 shows the cut-flow of the mistimed events.

Table A.1: Cut-flow of the mistimed events

Selections	2015	2016
All mistimed events	123	51
Pre-selections (without any lepton selections)	120	51
At least one Large-R jet	118	51
Large-R jet selection	118	51
Boson-tag LP (HP)	67 (13)	23 (6)
Photon selection	2 (1)	3 (1)
Event topology	0 (0)	3 (1)

No 2015 mistimed events passed the cut-flow. One 2016 events passed the cut-flow with the HP boson-tag. The passed event has $p_{\mathrm{T}}^{\mathrm{Jcalo}} = 800$ GeV, in the p_{T} range (800 GeV $< |p_{\mathrm{T}}^{\mathrm{J}}| <$ 850 GeV) there are 11 normal events which passed the boosted HP cut-flow. Therefore the mistimed events have little effect on the analysis, and they are not included in the analysis.

Events that caused the HLT to timeout (or crash) are stored in the debug stream. It is recommended to check these events in search analyses. We confirmed no debug stream data are included in our SRs and CRs.

Appendix B

Limit morphing

Linear assumption is used to extrapolate the upper limits for the intermediate mass signals in the limit plots Figure 10.11 and 10.9. The linear assumption is tested by a morphing technique considering acceptance × efficiency of the neighbouring mass points.

The m_{lvJ} binning is determined by signal resolution. (75 GeV at 500 GeV and 135 GeV at 1 TeV) There is no necessary to sample with smaller mass step than that, therefore we do not need to apply morphing in the high-mass region. It might be effective to apply morphing in the low mass region, in which we have signal mass points with 100 GeV mass steps. A test has been performed in the low-mass region producing a intermediate mass point between 600 GeV and 700 GeV in the HVT W′ model. Table B.1 summarises the results. There is close agreement (∼10 %) between the linear assumption and the morphing result, thus we do not need to apply morphing.

TABLE B.1: Comparison between linear assumption and morphing output. Each column shows the observed / expected / expected+1σ / expected+2σ/expected-1σ and expected-2σ upper limits.

Upper limit to the HVT W′ model (pb)	observed	expected	+1σ	+2σ	-1σ	-2σ
HVT W′ 600 GeV	0.17	0.32	0.44	0.61	0.23	0.18
HVT W′ 700 GeV	0.097	0.15	0.21	0.28	0.11	0.08
HVT W′ 650 GeV (Linear assumption)	0.13	0.24	0.33	0.45	0.17	0.13
HVT W′ 650 GeV (Morphing)	0.12	0.26	0.36	0.48	0.19	0.14
Ratio (Linear/Morphing - 1)	0.08	-0.08	-0.08	-0.06	-0.11	-0.07

Appendix C

Event display

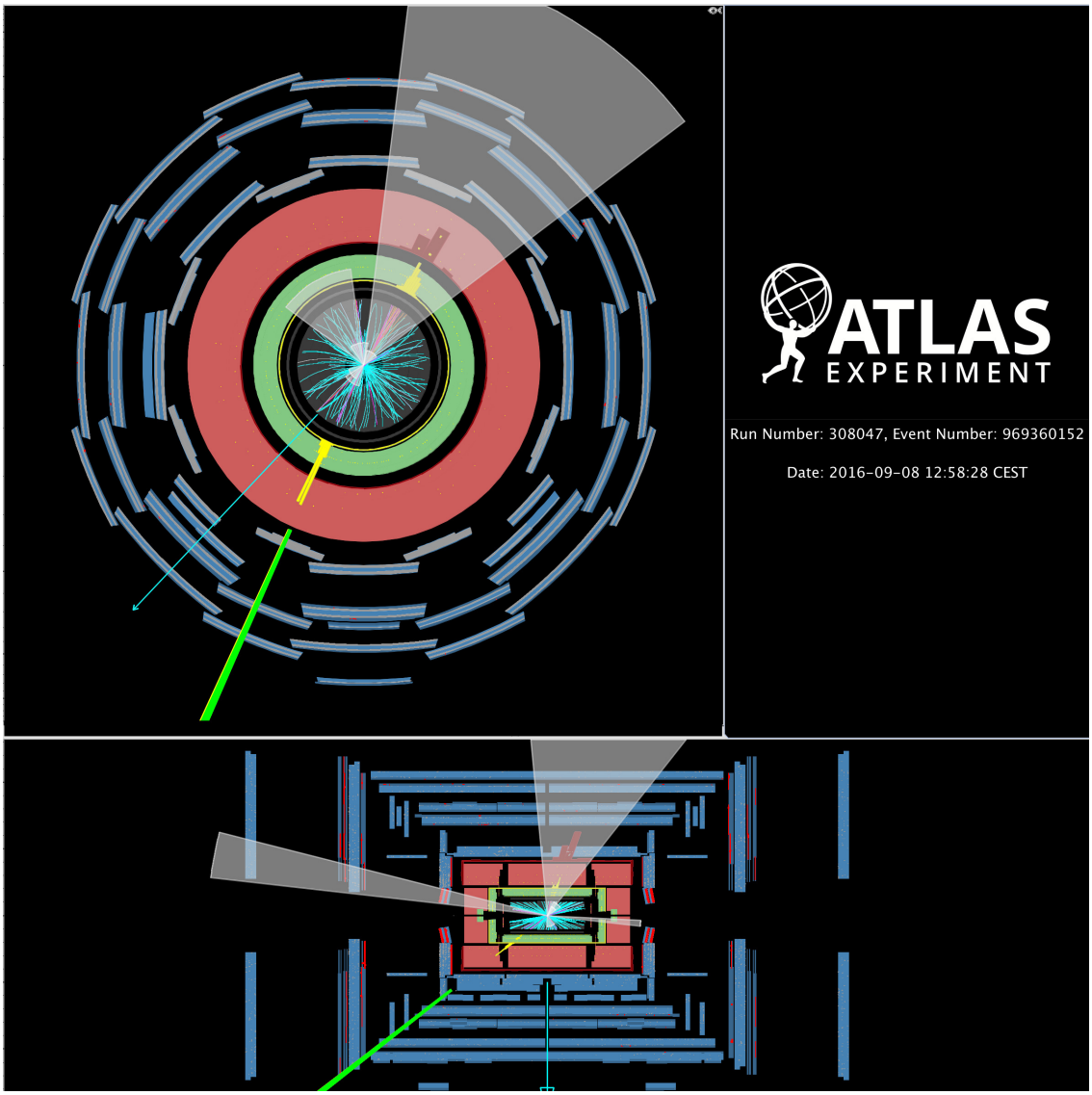


FIGURE C.1: Event display of the VBF candidate with the highest WV mass (2759 GeV) found in the merged low purity signal region. The hadronically decaying boson is the largest shaded cone with an 89 GeV mass, while the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction and the electron from the leptonic W decay are displayed as a blue arrow and a yellow tower in the calorimeters that is extrapolated to a green line outside the calorimeter volume, respectively. The leptonically decaying W boson is found to have p_{T} = 1138 GeV, as measured from combining the lepton with p_{T} = 777 GeV and $E_{\mathrm{T}}^{\mathrm{miss}}$ = 362 GeV, whereas the hadronically decaying V boson has p_{T} = 1118 GeV. The invariant mass of the tag-jets, appearing as small shaded cones in the forward regions, is measured to be 812 GeV while the pseudorapidity gap between them is 5.6 [38].

Bibliography

[1] G. Aad and et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) p. 1, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[2] S. Chatrchyan and et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) p. 30, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/S0370269312008581.

[3] F. Jegerlehner, *The hierarchy problem of the electroweak Standard Model revisited*, (2013), arXiv: 1305.6652 [hep-ph].

[4] Y. Akrami et al., *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, (2018), arXiv: 1807.06205 [astro-ph.CO].

[5] P. A. R. Ade et al., *Planck 2015 results. XIV. Dark energy and modified gravity*, Astron. Astrophys. **594** (2016) A14, arXiv: 1502.01590 [astro-ph.CO].

[6] G. R. Farrar and P. Fayet, *Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry*, Physics Letters B **76** (1978) p. 575, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/0370269378908584.

[7] P. Fayet, *Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions*, Physics Letters B **69** (1977) p. 489, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/0370269377908528.

[8] M. Dine and W. Fischler, *A Phenomenological Model of Particle Physics Based on Supersymmetry*, Phys. Lett. **110B** (1982) p. 227.

[9] M. J. Dugan, H. Georgi, and D. B. Kaplan, *Anatomy of a Composite Higgs Model*, Nucl. Phys. **B254** (1985) p. 299.

[10] K. Agashe, R. Contino, and A. Pomarol, *The Minimal composite Higgs model*, Nucl. Phys. B **719** (2005) p. 165, arXiv: hep-ph/0412089.

[11] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999) p. 3370, arXiv: hep-ph/9905221.

[12] L. Randall and R. Sundrum, *An Alternative to Compactification*, Phys. Rev. Lett. **83** (1999) p. 4690, arXiv: hep-th/9906064.

[13] H. Davoudiasl, J. Hewett, and T. Rizzo, *Experimental probes of localized gravity: On and off the wall*, Phys. Rev. D **63** (2001) p. 075004, arXiv: hep-ph/0006041.

[14] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 2016, arXiv: 1610.07922 [hep-ph].

[15] I. P. Ivanov, *Building and testing models with extended Higgs sectors*, Prog. Part. Nucl. Phys. **95** (2017) p. 160, arXiv: 1702.03776 [hep-ph].

[16] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, Phys. Rev. D **10** (1974) p. 275.

[17] H. Georgi and S. Glashow, *Unity of All Elementary Particle Forces*, Phys. Rev. Lett. **32** (1974) p. 438.

[18] H. Fritzsch and P. Minkowski, *Unified Interactions of Leptons and Hadrons*, Annals Phys. **93** (1975) p. 193.

[19] B. W. Lee, C. Quigg, and H. B. Thacker,
Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass,
Phys. Rev. **D16** (1977) p. 1519.

[20] B. W. Lee, C. Quigg, and H. B. Thacker,
The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass,
Phys. Rev. Lett. **38** (1977) p. 883.

[21] C. Collaboration, *Search for heavy resonances decaying to pairs of vector bosons in the l nu q qbar final state with the CMS detector in proton-proton collisions at sqrt s = 13 TeV*, (2017).

[22] "Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector", tech. rep. ATLAS-CONF-2018-030, CERN, 2018, URL: https://cds.cern.ch/record/2629411.

[23] A. M. Sirunyan et al., *Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV*,
Phys. Rev. Lett. **120** (2018) p. 081801, arXiv: 1709.05822 [hep-ex].

[24] G. Aad et al., *Measurements of $W^{\pm}Z$ production cross sections in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings*, Phys. Rev. **D93** (2016) p. 092004, arXiv: 1603.02151 [hep-ex].

[25] A. M. Sirunyan et al., *Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton–proton collisions at $\sqrt{s}$ = 13 TeV*,
Phys. Lett. **B774** (2017) p. 682, arXiv: 1708.02812 [hep-ex].

[26] J. R. Andersen et al., *Discovering Technicolor*, Eur. Phys. J. Plus **126** (2011) p. 81,
arXiv: 1104.1255 [hep-ph].

[27] A. Azatov, J. Galloway, and M. A. Luty, *Superconformal Technicolor: Models and Phenomenology*,
Phys. Rev. **D85** (2012) p. 015018, arXiv: 1106.4815 [hep-ph].

[28] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, *Heavy Vector Triplets: Bridging Theory and Data*, JHEP **09** (2014) p. 060, arXiv: 1402.4431 [hep-ph].

[29] J. de Blas, J. M. Lizana, and M. Perez-Victoria, *Combining searches of Z' and W' bosons*,
JHEP **01** (2013) p. 166, arXiv: 1211.2229 [hep-ph].

[30] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, *Warped Gravitons at the LHC and Beyond*,
Phys. Rev. D **76** (2007) p. 036006, arXiv: hep-ph/0701186.

[31] E. Witten, *Anti-de Sitter space and holography*, Adv. Theor. Math. Phys. **2** (1998) p. 253,
arXiv: hep-th/9802150 [hep-th].

[32] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo,
Phenomenology of the Randall-Sundrum Gauge Hierarchy Model, Phys. Rev. Lett. **84** (2000) p. 2080,
arXiv: hep-ph/9909255 [hep-ph].

[33] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov,
Gauge theory correlators from noncritical string theory, Phys. Lett. **B428** (1998) p. 105,
arXiv: hep-th/9802109 [hep-th].

[34] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) p. 063,
arXiv: 0802.1189 [hep-ph].

[35] "Performance of Top Quark and W Boson Tagging in Run 2 with ATLAS",
tech. rep. ATLAS-CONF-2017-064, CERN, 2017, URL: https://cds.cern.ch/record/2281054.

[36] ATLAS Collaboration, *Search for diboson resonances with boson-tagged jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, (2017), arXiv: 1708.04445 [hep-ex].

[37] M. Aaboud et al., *Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the $q\bar{q}^{(\prime)}b\bar{b}$ final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
Phys. Lett. **B774** (2017) p. 494, arXiv: 1707.06958 [hep-ex].

[38] M. Aaboud et al., *Search for WW/WZ resonance production in $\ell\nu qq$ final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **03** (2018) p. 042, arXiv: 1710.07235 [hep-ex].

[39] ATLAS Collaboration, *Searches for heavy ZZ and ZW resonances in the $\ell\ell qq$ and $\nu\nu qq$ final states in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, (2017), arXiv: 1708.09638 [hep-ex].

[40] A. M. Sirunyan et al., *Search for massive resonances decaying into WW, WZ, ZZ, qW, and qZ with dijet final states at sqrt(s) = 13 TeV*, (2017), arXiv: 1708.05379 [hep-ex].

[41] A. M. Sirunyan et al., *Search for heavy resonances that decay into a vector boson and a Higgs boson in hadronic final states at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. **C77** (2017) p. 636, arXiv: 1707.01303 [hep-ex].

[42] "Search for heavy resonances decaying to pairs of vector bosons in the l nu q qbar final state with the CMS detector in proton-proton collisions at sqrt s = 13 TeV", tech. rep. CMS-PAS-B2G-16-029, CERN, 2017, URL: https://cds.cern.ch/record/2296237.

[43] "Search for new heavy resonances decaying into a Z boson and a massive vector boson in the 2ℓ2q final state at $\sqrt{s}$ = 13 TeV", tech. rep. CMS-PAS-B2G-17-013, CERN, 2017, URL: https://cds.cern.ch/record/2296238.

[44] "Search for heavy resonances decaying into a Z boson and a vector boson in the $\nu\nu\, q\bar{q}$ final state", tech. rep. CMS-PAS-B2G-17-005, CERN, 2017, URL: https://cds.cern.ch/record/2273910.

[45] A. M. Sirunyan et al., *Search for diboson resonances in the $2\ell 2\nu$ final state*, (2017), arXiv: 1711.04370 [hep-ex].

[46] "Summary plots from the ATLAS Exotic physics group", 2017, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/.

[47] "CMS Beyond-two-generations (B2G) Public Physics Results - Dibosons", 2017, URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsB2GDibosons.

[48] "In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions", tech. rep. ATLAS-CONF-2017-063, CERN, 2017, URL: https://cds.cern.ch/record/2275655.

[49] "CERN web: Accelerators and Schedules", 2017, URL: https://beams.web.cern.ch/content/accelerators-and-schedules.

[50] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001, URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[51] M. Lamont, *Status of the LHC*, Journal of Physics: Conference Series **455** (2013) p. 012001, URL: http://stacks.iop.org/1742-6596/455/i=1/a=012001.

[52] R. Bruce et al., "LHC Run 2: Results and challenges", *Proceedings, 57th ICFA Advanced Beam Dynamics Workshop on High-Intensity and High-Brightness Hadron Beams (HB2016): Malmö, Sweden, July 3-8, 2016*, 2016 MOAM5P50, URL: http://inspirehep.net/record/1639591/files/moam5p50.pdf.

[53] "LHC Performance Workshop 2017", 2017, URL: https://indico.cern.ch/event/580313/.

[54] "LHC Performance Workshop 2018", 2018, URL: https://indico.cern.ch/event/676124/.

[55] H. Damerau, A. Findlay, S. Gilardoni, and S. Hancock, *RF Manipulations for Higher Brightness LHC-Type Beams*, (2013) 3 p, URL: https://cds.cern.ch/record/1595719.

[56] Y. Papaphilippou, H. Bartosik, G. Rumolo, and D. Manglunki, *Operational beams for the LHC*, (2014), arXiv: 1412.7857 [physics.acc-ph].

[57] S. Fartoukh et al., *Experimental validation of the Achromatic Telescopic Squeezing (ATS) scheme at the LHC*, J. Phys. Conf. Ser. **874** (2017) p. 012010, [,WEPIK030(2017)].

[58] B. Muratori and T. Pieloni, "Luminosity levelling techniques for the LHC", *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, [,177(2014)], 2014 p. 177, arXiv: 1410.5646 [physics.acc-ph].

[59] "ATLAS luminosity public results run-2", 2017,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

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

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

[62] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, PoS **EPS-HEP2015** (2015) p. 261,
arXiv: 1608.07850 [physics.ins-det].

[63] T. Lari, *Measurements of spatial resolution of ATLAS pixel detectors*,
Nucl. Instrum. Meth. **A465** (2000) p. 112, arXiv: physics/0012046 [physics.ins-det].

[64] G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*,
Eur. Phys. J. **C74** (2014) p. 3071, arXiv: 1407.5063 [hep-ex].

[65] M. Aaboud et al., *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77** (2017) p. 317,
arXiv: 1611.09661 [hep-ex].

[66] B. Carlson, T. M. Hong, and Atlas Collaboration,
"ATLAS level-1 calorimeter trigger: Run-2 performance and Phase-1 upgrades",
APS April Meeting Abstracts, 2017 S9.005.

[67] M. Sessa, *Performance of the ATLAS Level-1 muon barrel trigger during the Run 2 data taking*,
PoS **EPS-HEP2017** (2017) p. 803.

[68] N. Ellis et al., "The ATLAS level-1 muon to central trigger processor interface (MUCTPI)",
8th Workshop on Electronics for LHC Experiments Colmar, France, September 9-13, 2002, 2002 p. 227.

[69] "Data Quality Information for Data", 2017,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[70] "ATLAS Data Quality Information for Data", 2017,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[71] M. Aaboud et al.,
Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC,
Eur. Phys. J. **C76** (2016) p. 653, arXiv: 1608.03953 [hep-ex].

[72] S. van der Meer, "Calibration of the effective beam height in the ISR",
tech. rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, 1968,
URL: http://cds.cern.ch/record/296752.

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

[74] S. Moch, *Expectations at LHC from hard QCD*, J. Phys. **G35** (2008) p. 073001,
arXiv: 0803.0457 [hep-ph].

[75] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne,
Parton distributions in the LHC era: MMHT 2014 PDFs, Eur. Phys. J. **C75** (2015) p. 204,
arXiv: 1412.3989 [hep-ph].

[76] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) p. 026,
arXiv: hep-ph/0603175 [hep-ph].

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

[78] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al.,
Event generation with SHERPA 1.1, JHEP **0902** (2009) p. 007, arXiv: 0811.4622 [hep-ph].

[79] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) p. 639, arXiv: 0803.0883 [hep-ph].

[80] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) p. 079, arXiv: 1405.0301 [hep-ph].

[81] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) p. 040, arXiv: 1410.8849 [hep-ph].

[82] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) p. 039, arXiv: 0808.3674 [hep-ph].

[83] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) p. 111601, arXiv: 1111.5206 [hep-ph].

[84] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) p. 038, arXiv: 0709.1027 [hep-ph].

[85] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) p. 027, arXiv: 1207.5030 [hep-ph].

[86] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) p. 244, arXiv: 1207.1303 [hep-ph].

[87] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) p. 012, arXiv: hep-ph/0201195 [hep-ph].

[88] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys.Rev. **D82** (2010) p. 074018, arXiv: 1005.3457 [hep-ph].

[89] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. C **72** (2012) p. 2225, arXiv: 1206.0041 [hep-ph].

[90] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. D **69** (2004) p. 094008, arXiv: hep-ph/0312266.

[91] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) p. 252004, arXiv: 1303.6254 [hep-ph].

[92] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Compt/ Phys. Commun. (2014) p. 2930, arXiv: 1112.5675 [hep-ph].

[93] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) p. 054028, arXiv: 1001.5034 [hep-ph].

[94] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. **D82** (2010) p. 054018, arXiv: 1005.4451 [hep-ph].

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

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

[97] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **04** (2009) p. 002, arXiv: 0812.0578 [hep-ph].

[98] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02** (2010) p. 037, arXiv: 0911.5299 [hep-ph].

[99] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) p. 074024, arXiv: 1007.2241 [hep-ph].

[100] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **09** (2014) p. 145, arXiv: 1406.3660 [hep-ex].

[101] A. Hill and J. J. van der Bij, *Strongly Interacting Singlet - Doublet Higgs Model*, Phys. Rev. D **36** (1987) p. 3463.

[102] V. Barger, P. Langacker, M. McCaskey, M. J. Ramsey-Musolf, and G. Shaughnessy, *LHC Phenomenology of an Extended Standard Model with a Real Scalar Singlet*, Phys. Rev. D **77** (2008) p. 035005, arXiv: 0706.4311 [hep-ph].

[103] V. D. Barger, W.-Y. Keung, and E. Ma, *A Gauge Model With Light W and Z Bosons*, Phys. Rev. **D22** (1980) p. 727.

[104] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, *On the effect of resonances in composite Higgs phenomenology*, JHEP **10** (2011) p. 081, arXiv: 1109.1570 [hep-ph].

[105] B. Bellazzini, C. Csáki, and J. Serra, *Composite Higgses*, Eur. Phys. J. C **74** (2014) p. 2766, arXiv: 1401.2457 [hep-ph].

[106] G. Panico and A. Wulzer, *The Composite Nambu-Goldstone Higgs*, Lect. Notes Phys. **913** (2016) p. 1, arXiv: 1506.01961 [hep-ph].

[107] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[108] P. de Aquino, K. Hagiwara, Q. Li, and F. Maltoni, *Simulating graviton production at hadron colliders*, JHEP **06** (2011) p. 132, arXiv: 1101.5499 [hep-ph].

[109] T. Han, J. D. Lykken, and R.-J. Zhang, *On Kaluza-Klein states from large extra dimensions*, Phys. Rev. D **59** (1999) p. 105006, arXiv: hep-ph/9811350.

[110] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) p. 250.

[111] G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) p. 823, arXiv: 1005.4568 [physics.ins-det].

[112] M. Aaboud et al., *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. **C77** (2017) p. 673, arXiv: 1704.07983 [hep-ex].

[113] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) p. 444.

[114] F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, JINST **11** (2016) p. C12060.

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

[116] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, (2008).

[117] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, (2012).

[118] "Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV", tech. rep. ATL-PHYS-PUB-2016-015, CERN, 2016, URL: https://cds.cern.ch/record/2203514.

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

[120] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) p. 87, ISSN: 0734-189X, URL: http://www.sciencedirect.com/science/article/pii/S0734189X88800331.

[121] G. Aad et al.,
Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV,
Eur. Phys. J. **C76** (2016) p. 292, arXiv: 1603.05598 [hep-ex].

[122] G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*,
Eur. Phys. J. **C77** (2017) p. 490, arXiv: 1603.02934 [hep-ex].

[123] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber,
Longitudinally invariant K_t clustering algorithms for hadron hadron collisions,
Nucl. Phys. **B406** (1993) p. 187.

[124] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*,
Phys. Rev. **D48** (1993) p. 3160, arXiv: hep-ph/9305266 [hep-ph].

[125] T. A. collaboration, *Tagging and suppression of pileup jets*, (2014).

[126] "Optimisation of the ATLAS *b*-tagging performance for the 2016 LHC Run",
tech. rep. ATL-PHYS-PUB-2016-012, CERN, 2016, URL: https://cds.cern.ch/record/2160731.

[127] "Expected performance of the ATLAS *b*-tagging algorithms in Run-2",
tech. rep. ATL-PHYS-PUB-2015-022, CERN, 2015, URL: https://cds.cern.ch/record/2037697.

[128] G. Aad et al., *Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **09** (2013) p. 076, arXiv: 1306.4945 [hep-ex].

[129] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam,
Jet substructure as a new Higgs search channel at the LHC, Phys. Rev. Lett. **100** (2008) p. 242001,
arXiv: 0802.2470 [hep-ph].

[130] S. D. Ellis, C. K. Vermilion, and J. R. Walsh,
Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches,
Phys. Rev. **D81** (2010) p. 094023, arXiv: 0912.0033 [hep-ph].

[131] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, JHEP **02** (2010) p. 084,
arXiv: 0912.1342 [hep-ph].

[132] T. A. collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, (2016).

[133] A. J. Larkoski, I. Moult, and D. Neill, *Power Counting to Better Jet Observables*, JHEP **12** (2014) p. 009,
arXiv: 1409.6298 [hep-ph].

[134] M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. **D96** (2017) p. 072002,
arXiv: 1703.09665 [hep-ex].

[135] M. Aaboud et al., *In situ calibration of large-R jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector*, Submitted to: Eur. Phys. J. (2018), arXiv: 1807.09477 [hep-ex].

[136] ATLAS Collaboration,
Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data,
Eur. Phys. J. **C74** (2014) p. 3071, arXiv: 1407.5063 [hep-ex].

[137] ATLAS Collaboration,
Determination of the jet energy scale and resolution at ATLAS using Z/γ–jet events in data at $\sqrt{s}$ = 8 TeV,
ATLAS-CONF-2015-057 (2015), URL: https://cds.cern.ch/record/2059846.

[138] ATLAS Collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Phys. Rev. **D83** (2011) p. 052005, arXiv: 1012.4389 [hep-ex].

[139] M. Aaboud et al., *Measurement of the cross section for inclusive isolated-photon production in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector*, Phys. Lett. **B770** (2017) p. 473,
arXiv: 1701.06882 [hep-ex].

[140] ATLAS Collaboration, *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s}$ = 8 TeV*,
ATLAS-CONF-2015-017 (2015), URL: https://cds.cern.ch/record/2008678.

[141] ATLAS Collaboration,
Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ *TeV*,
Eur. Phys. J. **C73** (2013) p. 2304, arXiv: 1112.6426 [hep-ex].

[142] "Public Egamma Trigger Plots for Collision Data", 2017,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults.

[143] G. Aad et al., *Measurement of the polarisation of W bosons produced with large transverse momentum in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS experiment*, Eur. Phys. J. **C72** (2012) p. 2001,
arXiv: 1203.2165 [hep-ex].

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

[145] M. Aaboud et al., *Measurement of the Inelastic Proton-Proton Cross Section at* $\sqrt{s} = 13$*TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) p. 182002, arXiv: 1606.02625 [hep-ex].

[146] *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a* $b\bar{b}$ *pair in pp collisions at 13 TeV using the ATLAS detector*, ATLAS-CONF-2016-091 (2016),
URL: https://cds.cern.ch/record/2206813.

[147] D. Börner, *A method for the construction of strongly reduced representations of ATLAS experimental uncertainties and the application thereof to the jet energy scale*, EPJ Web Conf. **137** (2017) p. 11003.

[148] "Improving jet substructure performance in ATLAS using Track-CaloClusters",
tech. rep. ATL-PHYS-PUB-2017-015, CERN, 2017, URL: https://cds.cern.ch/record/2275636.

[149] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) p. 033006, arXiv: 1506.07443 [hep-ph].

[150] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) p. 023001,
arXiv: 1510.03865 [hep-ph].

[151] L. Moneta et al., "The RooStats Project", *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research, ACAT2010*, Proceedings of Science, 2010,
arXiv: 1009.1003.

[152] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra",
Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011, 2011,
arXiv: 1103.0354 [physics.data-an],
URL: https://inspirehep.net/record/891252/files/arXiv:1103.0354.pdf.

[153] G. Cowan, K. Cranmer, E. Gross, and O. Vitells,
Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71** (2011) p. 1554.

[154] A.L. Read, *Presentation of search results: the* CL_s *technique*, J. Phys. **G28** (2002) p. 2693.

[155] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke,
Interpolation between multi-dimensional histograms using a new non-linear moment morphing method,
Nucl. Instrum. Meth. **A771** (2015) p. 39, arXiv: 1410.7388 [physics.data-an].

[156] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*,
JHEP **11** (2001) p. 063, arXiv: hep-ph/0109231 [hep-ph].

[157] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*,
JHEP **05** (2009) p. 053, arXiv: 0903.1219 [hep-ph].

[158] M. Aaboud et al., *Combination of searches for heavy resonances decaying into bosonic and leptonic final states using 36* fb^{-1} *of proton-proton collision data at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*,
Phys. Rev. **D98** (2018) p. 052008, arXiv: 1808.02380 [hep-ex].

[159] A. Djouadi and J. Quevillon,
The MSSM Higgs sector at a high M_{SUSY}*: reopening the low tanβ regime and heavy Higgs searches*,
JHEP **10** (2013) p. 028, arXiv: 1304.1787 [hep-ph].

[160] M. Aaboud et al., *Search for anomalous electroweak production of WW/WZ in association with a high-mass dijet system in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, Phys. Rev. **D95** (2017) p. 032001, arXiv: 1609.05122 [hep-ex].

[161] "HL-LHC prospects for diboson resonance searches and electroweak vector boson scattering in the $WW/WZ \to \ell\nu qq$ final state", tech. rep. ATL-PHYS-PUB-2018-022, CERN, 2018, URL: https://cds.cern.ch/record/2645269.