

Università degli Studi di Napoli *Federico II*

DOTTORATO DI RICERCA IN  
FISICA

Ciclo XXXII

Coordinatore: Prof. Salvatore Capozziello

**Searches for new high mass resonances in  
the diboson channel with semileptonic  $llqq$   
final state in  $pp$  collisions at  $\sqrt{s} = 13$  TeV  
with the ATLAS detector**

Settore Scientifico Disciplinare FIS/01

**Dottorando**

Marco Lavorgna

**Tutori**

Prof. Leonardo Merola

Dott. Gianpaolo Carlino

Dott. Francesco Alessandro Conventi

ANNI 2017/2020

# Contents

<b>1</b>	<b>The Standard Model and Physics Beyond the SM</b>	<b>6</b>
1.1	Elements of quantum field theory . . . . .	7
1.1.1	Quantum electrodynamics . . . . .	8
1.1.2	Non abelian gauge group . . . . .	11
1.1.3	Quantum chromodynamics . . . . .	13
1.2	The Standard Model . . . . .	14
1.2.1	Spontaneous symmetry breaking . . . . .	14
1.2.2	Fermion sector . . . . .	18
1.3	Physics Beyond the Standard Model . . . . .	20
1.3.1	Two-Higgs Doublet Model . . . . .	21
1.3.2	Heavy Vector Triplet Model . . . . .	22
1.3.3	Randall–Sundrum Graviton . . . . .	23
1.3.4	Randall–Sundrum Radion . . . . .	23
1.3.5	Benchmark models for resonant diboson production . . . . .	24
<b>2</b>	<b>The ATLAS experiment at the LHC</b>	<b>28</b>
2.1	The LHC storage ring . . . . .	28
2.1.1	Technical specifications . . . . .	29
2.1.2	Data-taking . . . . .	33
2.2	The ATLAS experiment . . . . .	34
2.2.1	ATLAS Coordinate System . . . . .	37
2.3	ATLAS experimental apparatus . . . . .	39
2.3.1	Magnetic system . . . . .	39
2.3.2	Inner detector . . . . .	40
2.3.3	Calorimetric system . . . . .	42
2.3.4	Muon Spectrometer . . . . .	49
2.3.5	The ATLAS trigger system . . . . .	51
2.3.6	The Run-3 and the High Luminosity LHC upgrades. . . . .	54

<b>3</b>	<b>Physics objects reconstruction in ATLAS</b>	<b>56</b>
3.1	Electrons reconstruction and identification . . . . .	56
3.2	Muons reconstruction and identification . . . . .	60
3.3	Jets reconstruction . . . . .	64
3.4	b-tagging . . . . .	74
3.5	Missing energy reconstruction . . . . .	78
<b>4</b>	<b>Search for heavy-diboson resonances in the <math>llqq</math> final state</b>	<b>82</b>
4.1	Data and Monte Carlo samples . . . . .	83
4.2	Object reconstruction . . . . .	85
4.2.1	Leptons . . . . .	85
4.2.2	Jets . . . . .	90
4.2.3	Overlap Removal . . . . .	95
4.3	The event selection . . . . .	95
4.3.1	Trigger requirements . . . . .	97
4.3.2	$Z \rightarrow ll$ selection . . . . .	98
4.3.3	The production mechanism categorization . . . . .	101
4.3.4	$V \rightarrow J/ V \rightarrow qq$ selection . . . . .	105
4.4	Background estimation . . . . .	112
4.5	Systematic uncertainties . . . . .	124
4.5.1	Experimental uncertainties . . . . .	124
4.5.2	Background modeling uncertainties . . . . .	126
4.5.3	Signal theoretical uncertainties . . . . .	127
4.6	Statistical approach and results . . . . .	127
4.6.1	Likelihood function and test statistic . . . . .	130
4.6.2	Results . . . . .	133
<b>A</b>	<b>Standard Model production cross sections</b>	<b>149</b>
	<b>Bibliography</b>	<b>154</b>

# Introduction

The Standard Model (SM) is a non abelian gauge theory that describes the interactions between all known elementary particles.

The first step for the construction of the Standard Model was the development of a quantum field theory for the electromagnetic interactions (QED). Then in the 60s weak and electromagnetic interactions have been unified thanks to the Yang-Mills theories for non abelian gauge symmetries, through the model of Glashow-Weinberg-Salam (GWS). Finally the electroweak theory has been unified with the quantum chromodynamics (QCD).

Standard Model predictions have been verified with high precision by different experiments over the years after its formulation. The last fundamental particle foreseen in the SM was the Higgs boson, the quantum excitation of the field responsible for the electroweak symmetry breaking and the generation of particles masses. It was discovered in 2012, at CERN, where Higgs events were produced in proton-proton collisions at the Large Hadron Collider (LHC).

Despite all these predictions, there are many aspects that make this model not completely satisfactory. As an example, the Standard Model is unable to explain the matter-antimatter asymmetry in the universe or the neutrino oscillations, that imply that neutrinos have mass despite the Standard Model prediction. Moreover the interactions of the elementary particles are described at the *electroweak scale* ( $\mathcal{O}(100)$  GeV), while the energy scale to which gravitational effects can't be neglected is the *Planck scale* ( $10^{19}$  GeV), and no explanation has been found for this energy scale difference.

All these aspects suggest that the Standard Model represents only a good approx of a more general theory that incorporates also the gravitational interactions.

Many theoretical models of “*Beyond Standard Model Physics*” (BSM) have been developed and a major goal of the physics programme at the Large Hadron Collider (LHC) is the search for new phenomena that may become visible in high-energy proton-proton (pp) collisions.

One possible signature for new physics is the production of a heavy resonance

with the subsequent decay into a pair of vector bosons ( $WW$ ,  $WZ$ ,  $ZZ$ ). Several models of physics beyond the Standard Model predict such a signature: the Heavy Vector Triplet (HVT) phenomenological Lagrangian model, where a new heavy vector triplet ( $W'$ ,  $Z'$ ) is introduced with the new gauge bosons degenerate in mass; warped extra dimensions Randall–Sundrum (RS) models predict spin-2 Kaluza–Klein (KK) excitations of the graviton,  $G_{KK}$ , and spin-0 Radions.

Searches in final states in which one boson decays in leptons and the other one decays in hadrons (semileptonic analyses) are reported in this thesis work. Previous searches in these final states have been performed, using the statistics recorded during the 2015 and the beginning of 2016 by the ATLAS experiment at  $\sqrt{s} = 13$  TeV.

The integrated luminosity of the current work is  $139 \text{ fb}^{-1}$ , recorded in 2015–2018 at  $\sqrt{s} = 13$  TeV, corresponding to the full Run-2 statistics. The results of this work are going to be published in summer 2020.

The semileptonic searches lead to different research analyses:  $X \rightarrow ZV \rightarrow llqq$ ,  $X \rightarrow ZV \rightarrow \nu\nu qq$ ,  $X \rightarrow WV \rightarrow l\nu qq$ , where  $X$  represents a new resonance,  $V$  the  $W$  or  $Z$  vector boson,  $l$  and  $\nu$  are the leptons and the neutrinos coming from the leptonic decays of the vector bosons and  $q$  represents the quarks coming from the hadronic decays of the vector bosons. Production through gluon–gluon fusion (ggF), Drell–Yan (DY) and vector-boson fusion (VBF) processes are considered, depending on the assumed model. Heavy resonances would manifest themselves as resonant structures above the SM background in the invariant-mass distributions of the final state and in the  $X \rightarrow ZV \rightarrow \nu\nu qq$  channel as broad enhancements in the transverse-mass distributions of the  $\nu\nu qq$  final state.

My work has been dedicated on the  $X \rightarrow ZV \rightarrow llqq$  research channel. The thesis is organized in 4 chapters:

- in chapter 1 an overview of the Standard Model and of the fundamental interactions is reported, together with a description of the models of Physics beyond the Standard Model that predict the existence of a heavy resonance that decays in a pair of vector bosons;
- in chapter 2 the characteristics of the Large Hadron Collider and of the ATLAS experiment are presented;
- in chapter 3 the reconstruction and identification of the physics objects used in the analysis are described;
- chapter 4 presents the analysis performed in the  $X \rightarrow ZV \rightarrow llqq$  research channel. A detailed description of all the analysis features is

reported and the final results obtained are also described in this chapter. No significant deviations from expected background contributions have been observed, and constraints on the production cross sections of heavy resonances times the branching ratios in the diboson couple have been derived for all the tested signal models.

# Chapter 1

## The Standard Model and Physics Beyond the SM

The *Standard Model (SM)* of particle physics is a Quantum Field Theory (QFT), that describes all known elementary particles as well as their interactions, which are governed by three of the four known fundamental forces: strong force, electromagnetic force, and weak force [1, 2, 3, 4, 5, 6].

The particles, in the framework of the SM, are classified in two groups according to their spin:

- *Fermions* are half-integer spin particles (obeying Fermi-Dirac statistics) that form the matter particles. The fermions are categorized in leptons and quarks. The former interact only by the electromagnetic and weak forces, while the quarks interact also strongly and so they carry a color charge (in addition to an electric charge). Each fermion has a corresponding anti-particle, characterized by the same particle mass but opposite quantum numbers.
- *Bosons* are integer spin particles (obeying Bose-Einstein statistics) that are responsible of interactions among fermions. In the the electromagnetism, the exchange of a *virtual* photon mediates the interactions between charged particles. The strong force is mediated by eight massless gluons; at large-scale it is responsible for bounding together protons and neutrons in the atomic nucleus. The weak force is responsible for the nuclear  $\beta$ -decays of certain radioactive isotopes and the nuclear fusion processes of the Sun and it is mediated by the exchange of  $W^\pm$  and  $Z$  bosons.

All the fundamental forces, with the exception of gravity, are described by a quantum field theory. Figure 1.1 shows a map of all the elementary

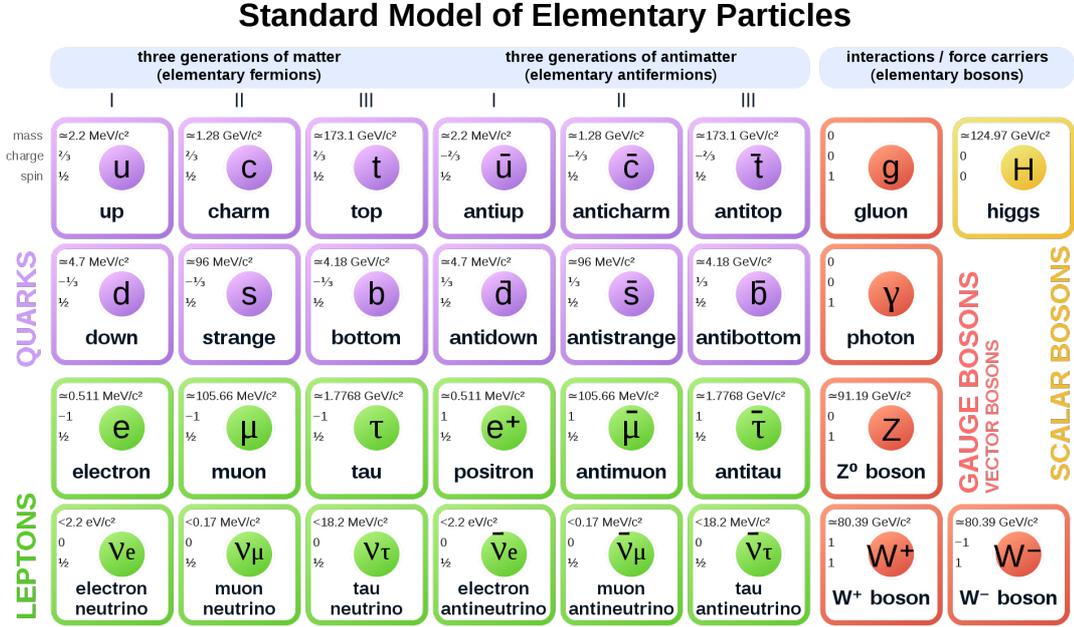


Figure 1.1: The Standard Model of elementary particles.

particles foreseen by the SM and observed.

## 1.1 Elements of quantum field theory

The quantum field theory is a physics theory that describes the behaviour of subatomic particles and their interactions, incorporating quantum mechanics, special relativity and the concept of field.

The particles are described as field operators. The starting point is indeed the so called *second quantization* of the fields, which describes the field operators as a linear superimposition of operators who, when applied to state vectors, create or destroy a particle. The field operators, functions of the space-time coordinates, obey to equations of motion derived from a lagrangian  $L$  through a variational principle. Usually  $L$  is written as integral on the spatial coordinates of a lagrangian density  $\mathcal{L}$ , simply referred as lagrangian:

$$L(t) = \int d^3\mathbf{x} \mathcal{L}(x).$$

$\mathcal{L}$  is function of the  $\phi_j(x)$  fields and of their gradients  $\frac{\partial\phi_j(x)}{\partial x_\mu} \equiv \partial_\mu\phi_j(x)$ :

$$\mathcal{L}(t, \mathbf{x}) = \mathcal{L}(\phi_j(x), \partial_\mu \phi_j(x)). \quad (1.1)$$

The integral of  $\mathcal{L}(x)$  on the space and the time gives the action  $S$ :

$$S = \int L(t)dt = \int d^4x \mathcal{L}(\phi_j(x), \partial_\mu \phi_j(x)).$$

The equations of motion of the fields are derived by the Eulero-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial \phi_j} - \frac{\partial}{\partial x_\mu} \left( \frac{\partial \mathcal{L}}{\partial (\frac{\partial \phi_j}{\partial x_\mu})} \right) = 0 \quad (j = 1, 2, \dots) \quad (1.2)$$

that come from the variational principle:

$$\delta \int_{t_1}^{t_2} L(t)dt = 0$$

where the times  $t_1$  and  $t_2$  are arbitrary and the variations of the fields for  $t = t_1$  and  $t = t_2$  are assumed to be null.  $\mathcal{L}$  needs to be a Lorentz scalar in order to have equations of motion that are covariant.

Interactions between the fields are introduced by imposing that the lagrangian  $\mathcal{L}_0$ , which describes the motion of free fields, satisfies a *local gauge symmetry*. In this way one get the interaction term  $\mathcal{L}'$ , so that:

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}'.$$

### 1.1.1 Quantum electrodynamics

Quantum electrodynamics (QED) is the quantum theory that describes the electromagnetic interactions between charged particles. QED is an abelian gauge theory, with the symmetry group being  $U(1)$ .

The lagrangian of a free electron [7] is:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (1.3)$$

This lagrangian is invariant under a *global gauge transformation*, defined as:

$$\psi \longrightarrow \psi' = e^{iq\theta}\psi$$

where  $q$  is the electron charge and  $\theta$  is an arbitrary constant. Because a finite transformation can be built by an infinite series of infinitesimal transformations, the lagrangian of a free electron is invariant also under infinitesimal transformations, like:

$$\psi \longrightarrow \psi' = (1 + iq\epsilon)\psi \quad (1.4)$$

where  $\epsilon$  is an arbitrary infinitesimal constant. The infinitesimal transformation 1.4 gives infinitesimal variations:

$$\delta\psi = \psi' - \psi = iq\epsilon\psi$$

$$\delta(\partial_\mu\psi) = \partial_\mu\psi' - \partial_\mu\psi = iq\epsilon(\partial_\mu\psi).$$

Similar relations are obtained for  $\bar{\psi}$ . In correspondence to these variations, the variation of the lagrangian,  $\delta\mathcal{L}$ , is zero; Therefore one has:

$$2\epsilon\partial_\mu(q\bar{\psi}\gamma^\mu\psi) = 0$$

and from the arbitrariness of the infinitesimal constant  $\epsilon$  one has:

$$\partial_\mu j^\mu = \partial_\mu(q\bar{\psi}\gamma^\mu\psi) = 0. \quad (1.5)$$

Global gauge invariance gives a *conserved current*, that in this case is the electromagnetic current. Considering that:

$$q = \int j^0 d^3\mathbf{x}$$

and that the equation 1.5 implies that  $\partial_0 j^0 = -\partial_k j^k$  with  $k = (1, 2, 3)$ , one has:

$$\frac{dq}{dt} = - \int_V \partial_i j^i d^3\mathbf{x} = - \int_S \mathbf{j} \cdot d\mathbf{s}. \quad (1.6)$$

Assuming that the fields, and so  $\mathbf{j}$ , go to zero at infinity sufficiently rapidly, for  $S \rightarrow \infty$  one has that the temporal derivative of the charge is zero. The global gauge invariance leads to the *conservation of the electric charge*. In this case the gauge transformations form a group,  $U(1)$ , abelian since the product between two transformations is commutative:

$$[U(\theta_1), U(\theta_2)] = 0.$$

A *local gauge transformation*, in  $U(1)$ , can be written as:

$$\psi \longrightarrow \psi' = e^{iq\theta(x)}\psi$$

where, unlike the global case,  $\theta(x)$  is a function of the coordinates. In the lagrangian 1.3, the mass term is invariant under local gauge transformations, while the field derivative is not:

$$\partial_\mu\psi \rightarrow \partial_\mu\psi' = e^{iq\theta(x)}\partial_\mu\psi(x) + iqe^{iq\theta(x)}\psi(x)\partial_\mu\theta(x)$$

with the last term destroying the invariance. To restore the invariance, it is necessary to introduce a derivative operator, that has the same phase transformation as the field:

$$\mathcal{D}_\mu\psi \rightarrow e^{iq\theta(x)}\mathcal{D}_\mu\psi.$$

This derivative is called *covariant derivative* and it's defined as:

$$\mathcal{D}_\mu \equiv \partial_\mu + iqA_\mu(x).$$

The invariance is satisfied as long as the vectorial field  $A_\mu(x)$  transforms as:

$$A_\mu(x) \rightarrow A_\mu(x) - \partial_\mu\theta(x). \quad (1.7)$$

The eq. 1.7 expresses the gauge liberty under the choice of the classical potential vector of the electromagnetic field.

Replacing in the lagrangian 1.3 the covariant derivative one has the invariant lagrangian under local gauge transformation:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\mathcal{D}_\mu\psi - m\bar{\psi}\psi.$$

Writing explicitly the covariant derivative:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma^\mu\psi A_\mu = \mathcal{L}_0 - j^\mu A_\mu.$$

The current  $j^\mu = q\bar{\psi}\gamma^\mu\psi$  has the form required by the global gauge invariance and the term  $-j^\mu A_\mu$  represents the interaction between the Dirac particle and the electromagnetic field.

The request of local gauge invariance led to the introduction of the gauge field  $A_\mu$ , that one associates with the physical photon field. In order to complete the QED lagrangian it is necessary to add the kinetic energy term  $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ , that describes the propagation of free photons and it's invariant under local gauge transformations:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - j^\mu A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}.$$

If the photon had mass, it would have been necessary to add a term of mass of the form  $(1/2)m_\gamma^2 A_\mu A^\mu$ , and this would destroy the local gauge invariance.

As summary, while the global gauge invariance leads to the conservation of the electric charge, the request of local gauge invariance leads to the introduction of a vectorial field,  $A_\mu$ , whose gauge boson is the photon that has to have null mass (consistently with the infinite range of the electromagnetic interactions) so that the local invariance is preserved. Moreover, the request of local gauge invariance leads to specifying the form of the radiation-matter interaction, that in the case of the electron (a Dirac particle) is  $-j^\mu A_\mu$ , where  $j^\mu$  is the conserved current of the free electron lagrangian.

### 1.1.2 Non abelian gauge group

In order to describe the other fundamental interactions, more complex symmetries than the abelian one have to be taken into account. The simplest non-abelian group of transformations is  $SU(2)$ . The fields enter in the theory as multiplets:

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix}$$

which form the basis for a group representation.

Considering the case of two non-interacting fields of spin 1/2,  $\psi_1$  and  $\psi_2$ , that form a doublet for an  $SU(2)$  symmetry:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

the lagrangian is given by the sum of two Dirac lagrangians:

$$\mathcal{L} = i\bar{\psi}_1 \gamma^\mu \partial_\mu \psi_1 - m_1 \bar{\psi}_1 \psi_1 + i\bar{\psi}_2 \gamma^\mu \partial_\mu \psi_2 - m_2 \bar{\psi}_2 \psi_2$$

and it can be written in matrix form as:

$$\mathcal{L} = i\bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi$$

where  $m$  is a mass diagonal matrix:

$$m = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix}$$

It is possible to show that  $\mathcal{L}$  is not invariant under an infinitesimal local gauge transformation:

$$\psi \longrightarrow \psi' = (1 - ig\boldsymbol{\epsilon}(x) \cdot \frac{\boldsymbol{\tau}}{2})\psi$$

where  $\boldsymbol{\tau}/2 \equiv (\tau_1/2, \tau_2/2, \tau_3/2)$  is the “isospin” operator, whose components are the generators of the transformation, and  $\boldsymbol{\epsilon}(x)$  is an arbitrary infinitesimal vector in the isospin space. The  $\tau_i$  matrices, with  $i = 1, 2, 3$  are the *Pauli matrices*.

In the same way of the abelian case, it is possible to define a covariant derivative in order to make the lagrangian invariant under the  $SU(2)$  symmetry transformation:

$$\mathcal{D}_\mu = \partial_\mu + \frac{i}{2}g\boldsymbol{\tau} \cdot \mathbf{W}_\mu$$

where the three gauge fields  $\mathbf{W}_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)$ , called *Yang-Mills gauge fields*, transform as:

$$\mathbf{W}_\mu(x) \longrightarrow \mathbf{W}_\mu(x) + \partial_\mu\boldsymbol{\epsilon}(x) + g\boldsymbol{\epsilon}(x) \times \mathbf{W}_\mu(x)$$

to maintain the gauge invariance. By the introduction of a kinetic energy term of the gauge fields,

$$\mathcal{L}_W = -\frac{1}{4}\mathbf{G}_{\mu\nu}\mathbf{G}^{\mu\nu}$$

one has the lagrangian given by the Yang-Mills theory [8]. In order to preserve the local gauge invariance, the tensor  $\mathbf{G}_{\mu\nu}$  is required to have the following form:

$$\mathbf{G}_{\mu\nu} = \partial_\mu\mathbf{W}_\nu - \partial_\nu\mathbf{W}_\mu - g\mathbf{W}_\mu \times \mathbf{W}_\nu.$$

Since the  $SU(2)$  group is non-commutative, self couplings of the gauge fields appear as shown by the term  $\mathbf{W}_\mu \times \mathbf{W}_\nu$ .

As in the abelian case, the gauge fields introduced have to be massless in order to preserve the gauge invariance. Indeed a mass term of the form  $m^2\mathbf{W}_\nu\mathbf{W}^\nu$  would violate the local invariance.

There is a biunivocal correspondence between the number of generators of the group symmetry and the number of gauge fields necessary to preserve the invariance of the lagrangian under such symmetry.

### 1.1.3 Quantum chromodynamics

Quantum chromodynamics (QCD) [9] is the quantum field theory for strong interactions between hadronic particles. It is based on a gauge invariance principle, with the symmetry group being the non-abelian colour  $SU(3)_C$ . It has been experimentally proven that for each flavour of quarks there are three different states. Each state is associated with a charge called “color”. So in the QCD description quarks participate in a color triplet, for each flavour  $f = u, d, c, s, t, b$ :

$$\psi^f(x) = \begin{pmatrix} \psi_1^f(x) \\ \psi_2^f(x) \\ \psi_3^f(x) \end{pmatrix} \quad (1.8)$$

The QCD free Lagrangian is given by:

$$\mathcal{L} = \sum_{f=1}^6 \overline{\psi^f} (i\gamma^\mu \partial_\mu - m) \psi^f. \quad (1.9)$$

An infinitesimal local gauge transformation in  $SU(3)_C$  has the form:

$$\psi^f(x) \rightarrow \psi'^f(x) = [1 + ig_s \theta_\alpha(x) T_\alpha] \psi^f(x) \quad (1.10)$$

where  $g_s = \sqrt{4\pi\alpha_s}$  is the strong coupling constant and  $T_\alpha$  ( $\alpha = 1, \dots, 8$ ) are the eight generators of the  $SU(3)_C$  group. The generators are conventionally taken as  $T_\alpha = \Lambda_\alpha/2$ , where  $\Lambda_\alpha$  are the eight  $3 \times 3$  *Gell-Mann matrices*. In order to make the lagrangian invariant under the  $SU(3)_C$  symmetry transformation, eight gauge gluon fields,  $A_\alpha(x)$ , are introduced and the following covariant derivative is defined:

$$\mathcal{D}_\mu = \partial_\mu + ig_s T_\alpha A_{\alpha\mu}(x). \quad (1.11)$$

The gluon fields have to transform, in order to preserve the local gauge invariance, as:

$$A_{\alpha\mu}(x) \rightarrow A'_{\alpha\mu}(x) = A_{\alpha\mu}(x) - \partial_\mu \theta_\alpha(x) - g_s f_{\alpha\beta\gamma} \theta_\beta(x) A_{\gamma\mu}(x)$$

where  $f_{\alpha\beta\gamma}$  are the *constants structure* of the  $SU(3)_C$  group.

In order to complete the full QCD lagrangian, the kinetic term,  $-\frac{1}{4}G_{\alpha\mu\nu}G^{\alpha\mu\nu}$ , needs to be added for each gluon. The tensor  $G_{\alpha\mu\nu}$  has the following form:

$$G_{\alpha\mu\nu} = \partial_\mu A_{\alpha\nu} - \partial_\nu A_{\alpha\mu} - g_s f_{\alpha\beta\gamma} A_{\beta\mu} A_{\gamma\nu}. \quad (1.12)$$

The complete QCD lagrangian is:

$$\mathcal{L} = \sum_f [\bar{\psi}^f (i\gamma^\mu \partial_\mu - m)\psi^f - g_s \bar{\psi}^f \gamma^\mu T_\alpha A_{\alpha\mu}(x)\psi^f] - \frac{1}{4} G_{\alpha\mu\nu} G^{\alpha\mu\nu} \quad (1.13)$$

where the index  $f = 1, \dots, 6$  indicates the quark flavour and the index  $\alpha = 1, \dots, 8$  the gluon field. The non-abelian property allows to gluons auto-interacting terms present in the above QCD lagrangian. These vertex are not present in the QED and the main consequences are the asymptotic freedom and colour confinement phenomena.

## 1.2 The Standard Model

The Standard Model is a gauge theory based on the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  symmetry group.

$SU(3)_C$  is the non-abelian colour symmetry group, for the description of the strong interactions between quarks. The weak and electromagnetic interactions have been unified through the model of Glashow-Weinberg-Salam (GWS) [10, 11, 12]. Electroweak interactions are described by the  $SU(2)_L \times U(1)_Y$  symmetry group, where the index  $L$  describes the chirality and the index  $Y$  the hypercharge [13].

The gauge group uniquely determines the interactions and the number of gauge vector bosons corresponding to the generators of the group (tab. 1.1). These are:

- Eight massless gluons, that correspond to the eight generators of  $SU(3)_C$ .
- Four gauge bosons, generators of the  $SU(2)_L \times U(1)_Y$  group. Three of these bosons are expected to be massive ( $W^\pm, Z^0$ ), and one massless ( $\gamma$ ). The three gauge bosons of the  $SU(2)$  group are usually indicated as  $W_\mu^1, W_\mu^2, W_\mu^3$ , while  $B_\mu$  is the boson associated to the  $U(1)$  group.

### 1.2.1 Spontaneous symmetry breaking

As discussed in the previous paragraph, the application of the gauge principle leads to massless vector bosons. With the exception of the photon and gluons, however, the other  $W^\pm$  and  $Z^0$  bosons are expected to be massive. The *Higgs mechanism* [15] gives a way to solve this inconsistency. In its minimal form,

Interaction	Boson	Charge[ $Q/e$ ]	Mass [GeV]
Electromagnetic	$\gamma$	0	0
Weak	$W^\pm$	$\pm 1$	$80.379 \pm 0.012$ GeV
	$Z$	0	$91.1876 \pm 0.0021$ GeV
Strong	$g$	0	0

Table 1.1: The gauge bosons of the three fundamental interactions described in the Standard Model [14].

this mechanism provides for the existence of a scalar boson which, when coupled with field bosons, allows the  $W^\pm$  and  $Z^0$  bosons to have a mass different from zero. Considering therefore a lagrangian of the following form:

$$\begin{aligned}
\mathcal{L}_{SU(2)_L \times U(1)_Y} &= \bar{\psi}_L i\gamma^\mu \mathcal{D}_\mu^L \psi_L + \bar{\psi}_R i\gamma^\mu \mathcal{D}_\mu^R \psi_R - \frac{1}{4} \mathbf{G}_{\mu\nu} \mathbf{G}^{\mu\nu} - \frac{1}{4} C_{\mu\nu} C^{\mu\nu} \\
&= \bar{\psi}_L \gamma^\mu \left[ i\partial_\mu - g \frac{\boldsymbol{\tau}}{2} \cdot \mathbf{W}_\mu - g' \frac{Y}{2} B_\mu \right] \psi_L + \\
&\quad + \bar{\psi}_R \gamma^\mu \left[ i\partial_\mu - g' \frac{Y}{2} B_\mu \right] \psi_R - \frac{1}{4} \mathbf{G}_{\mu\nu} \mathbf{G}^{\mu\nu} - \frac{1}{4} C_{\mu\nu} C^{\mu\nu}
\end{aligned} \tag{1.14}$$

where  $\psi_L$  and  $\psi_R$  represent a generic fermionic doublet and singlet and  $-1/4 \mathbf{G}_{\mu\nu} \mathbf{G}^{\mu\nu}$ ,  $-1/4 C_{\mu\nu} C^{\mu\nu}$  the kinetic energy terms of the gauge fields  $\mathbf{W}$  and  $B$ , in order to give mass to the  $\mathbf{W}$  fields, the *gauge symmetry breaking* is introduced.

To understand how the gauge symmetry breaking works, an isospin doublet of scalar complex fields, with hypercharge  $Y = 1$  (*Higgs doublet*) can be considered:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.15}$$

The four real scalar fields must be introduced in the form of a multiplet of  $SU(2)_L \times U(1)_Y$  to ensure that the invariance of the lagrangian remains. The lagrangian of the doublet, after the request of local gauge invariance, is:

$$\mathcal{L}_\phi = (\mathcal{D}_\mu^L \phi)^\dagger (\mathcal{D}^{\mu,L} \phi) - \mu^2 \phi^\dagger \phi - \frac{\lambda}{4} (\phi^\dagger \phi)^2 - \frac{1}{4} \mathbf{G}_{\mu\nu} \mathbf{G}^{\mu\nu} - \frac{1}{4} C_{\mu\nu} C^{\mu\nu} \tag{1.16}$$

where the kinetic terms are added from the equation 1.14. Taking into account the case where the potential  $V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{4} (\phi^\dagger \phi)^2$  has  $\mu^2 < 0$  and  $\lambda > 0$ , this potential is minima for:

$$\phi^\dagger\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}$$

and the set of points so described is invariant under  $SU(2)$  transformations. It is convenient to consider the choice:

$$\phi_1 = \phi_2 = \phi_4 = 0, \quad \phi_3 = v^2.$$

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad \text{with} \quad v = \sqrt{-\frac{\mu^2}{\lambda}}. \quad (1.17)$$

The vacuum state, given by 1.17, has quantum numbers  $I = 1/2$ ,  $I_3 = -1/2$  and  $Y = 1$ . The arbitrary choice of the vacuum state, among the equivalent possible minima, breaks spontaneously the  $SU(2)_L \times U(1)_Y$  symmetry. This choice any case leaves unchanged the electromagnetic gauge  $U(1)_{em}$  symmetry, having it as generator the operator of charge  $Q$ . Developing  $\phi(x)$  around the vacuum state  $\phi_0$ , one gets the following parametrization:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} e^{i\theta(x)\cdot\tau/2v}. \quad (1.18)$$

Using a suitable gauge transformation, the Higgs doublet is reduced to:

$$\phi'(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \quad (1.19)$$

Replacing this expression in the lagrangian 1.16, one has:

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\partial_\mu H\partial^\mu H - \mu^2 H^2 \\ & - \frac{1}{4}(\partial_\mu W_{1\nu} - \partial_\nu W_{1\mu})(\partial^\mu W_1^\nu - \partial^\nu W_1^\mu) + \frac{1}{8}g^2 v^2 W_{1\mu}W_1^\mu \\ & - \frac{1}{4}(\partial_\mu W_{2\nu} - \partial_\nu W_{2\mu})(\partial^\mu W_2^\nu - \partial^\nu W_2^\mu) + \frac{1}{8}g^2 v^2 W_{2\mu}W_2^\mu \\ & - \frac{1}{4}(\partial_\mu W_{3\nu} - \partial_\nu W_{3\mu})(\partial^\mu W_3^\nu - \partial^\nu W_3^\mu) - \frac{1}{4}C_{\mu\nu}C^{\mu\nu} \\ & + \frac{1}{8}v^2(gW_{3\mu} - g'B_\mu)(gW_3^\mu - g'B^\mu). \end{aligned} \quad (1.20)$$

The first line is the lagrangian of a scalar field, the *Higgs field*, with mass:

$$m_H = \sqrt{2}|\mu| = \sqrt{2\lambda}v^2. \quad (1.21)$$

The following two lines show how the fields  $W_1^\mu$  and  $W_2^\mu$  of the triplet  $\mathbf{W}^\mu$  have acquired mass:

$$M_1 = M_2 = \frac{1}{2}gv \equiv M_W. \quad (1.22)$$

The last two lines show that the fields  $W_3^\mu$  and  $B^\mu$  are mixed. The last line moreover concerns the only combination  $gW_3^\mu - g'B^\mu$ , which evidently became massive. So it is possible to rearrange the last two lines replacing  $W_3^\mu$  and  $B^\mu$  fields with two new  $Z^\mu$  and  $A^\mu$  fields, obtained as orthogonal combinations of the previous ones.

The four fields of the electroweak theory can be written, in the end, as:

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \quad (1.23)$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \quad (1.24)$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2) \quad (1.25)$$

where  $W_\mu^\pm$  are the two vectorial bosons that mediate the weak interactions of charged current,  $Z_\mu$  is the vectorial boson that mediates the weak interactions of neutral current, and  $A_\mu$  is the photon, that mediates the electromagnetic interactions.

The Weinberg angle  $\theta_W$  is defined by the  $g$  and  $g'$  couplings of  $SU(2)_L$  and  $U(1)_Y$ :

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (1.26)$$

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \quad (1.27)$$

The electric charge can be written as:

$$e = g \sin\theta_W = g' \cos\theta_W \quad (1.28)$$

The hypothesis of the Higgs boson was raised in 1964, and phenomenological studies on its production mechanisms and decay began in the 1970s. Shortly afterwards were the first studies on the possibility of producing a Higgs boson from  $e^+e^-$ ,  $\bar{p}p$  e  $pp$  collisions. Before the start of the data taking at LHC, the most significant Higgs boson research was carried out at the CERN's Large Electron Positron collider (LEP) between 1989 and 2000, and

at the Tevatron at Fermilab. Subsequently, on July 4, 2012, the ATLAS and CMS experiments at CERN's LHC announced the discovery of a new particle candidate to be the Higgs boson with a mass of about 125 GeV [16, 17]. The identification of the new particle with the Higgs boson was supported by measurements of its properties, such as spin, parity and coupling properties, made during the first phase of LHC data collection at  $\sqrt{s} = 7-8$  TeV [18, 19].

### 1.2.2 Fermion sector

The quantum numbers of the electroweak theory are the weak isospin  $I$ , the hypercharge  $Y$  and the electric charge  $Q$ , linked by the Gell-Mann/Nishijima relation  $Q = I_3 + Y/2$ . Table 1.2 reports the quantum numbers for each fermion. Left-handed fermions exist as  $SU(2)_L$  doublet with  $I = 1/2$ , while right-handed fermions exist as singlet with  $I = 0$ . Right-handed neutrinos, as well as left-handed antineutrinos, have not been observed in nature so far. Both leptons and quarks are organised in three generations.

*Left-handed* fermions interact via  $W^\pm$  and  $Z^0$  bosons, while *right-handed* fermions interact via  $Z^0$  boson. The photon is electrically neutral and only pairs with charged particles.

In the Standard Model, electroweak interactions can be studied separately from strong interactions since the symmetry under the  $SU(3)_C$  color group is not broken and there is no mixing between the strong and electroweak sectors.

	Generations			Quantum numbers			
	1	2	3	$I$	$I_3$	$Y$	$Q[e]$
Lepton	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	1/2	1/2	-1	0
	$e^-_R$	$\mu^-_R$	$\tau^-_R$	1/2	-1/2	-1	-1
				0	0	-2	-1
Quark	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	1/2	1/2	1/3	2/3
	$u_R$	$c_R$	$t_R$	1/2	-1/2	1/3	-1/3
	$d_R$	$s_R$	$b_R$	0	0	4/3	2/3
				0	0	-2/3	1/3

Table 1.2: Overview of the fermion quantum numbers in the Standard Model.

The  $d'$ ,  $s'$ ,  $b'$  quark are eigenstate of the weak interaction obtained from the rotation of the mass eigenstates  $d$ ,  $s$ ,  $b$ , described through the CKM matrix (*Cabibbo-Kobayashi-Maskawa*):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \simeq \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.986 & 0.041 \\ 0.008 & 0.040 & 1.021 \end{pmatrix}$$

By emitting or absorbing a  $W^\pm$  boson, transitions can occur between different generations of quarks. The diagonal elements of the matrix are close to unit, while the off-diagonal elements refer to transitions between different families that are suppressed.

Fermion mass terms cannot appear directly in the Lagrangian because they would break the  $SU(2)_L \times U(1)_Y$  symmetry. Also fermions acquire mass interacting with the Higgs field. In a theory with spontaneously broken symmetry it is not necessary to explicitly introduce mass terms in the Lagrangian but it is possible to give mass to the fermions by coupling them to a scalar field, by means of a Yukawa like coupling:

$$\mathcal{L}_{\text{Yukawa}} = -G_\ell^{ij} \bar{L}_L^i \phi \ell_R^j - G_d^{ij} \bar{Q}_L^i \phi d_R^j - G_u^{ij} \bar{Q}_L^i \phi_C u_R^j + h.c. \quad (1.29)$$

where  $\bar{L}_L^i$  e  $\bar{Q}_L^i$  are the isospin doublet of the leptons and quarks and  $\ell_R^j$ ,  $d_R^j$ ,  $u_R^j$  are singlets for leptons and down and up quark states. In the third term:

$$\phi_C = i\sigma_2 \phi^* \quad (1.30)$$

is the conjugate doublet in the sense of  $SU(2)_L$  of the doublet  $\phi$ .

The matrices  $G_\ell^{ij}$ ,  $G_d^{ij}$  and  $G_u^{ij}$  define the coupling constant and the mixing of the quark generations. After spontaneous symmetry breaking,  $\phi$  takes the form in equation 1.19, and mass terms for the fermions are obtained.

The electron mass term is, for example:

$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}} v (\bar{e}_L e_R + \bar{e}_R e_L) - \frac{G_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) H = -m_e \bar{e} e - \frac{m_e}{v} \bar{e} e H \quad (1.31)$$

where  $m_e = G_e v / \sqrt{2}$  gives the electron mass.

The coupling constant  $G_e$  is arbitrary, so that the electron mass is not predicted by the theory. The Lagrangian has also an interaction term that couples the Higgs field to the electron field and this term is proportional to the electron mass.

The masses of the other fermions are generated with the same procedure, except neutrinos that assumed to be massless in the Standard Model.

The SM fermion sector depends on 13 independent parameters: 6 quark masses, 3 charged lepton masses (neutrinos are assumed to have zero mass), 3 mixing angles for quarks and a phase factor responsible for CP violation.

### 1.3 Physics Beyond the Standard Model

The experimental results, obtained at the accelerators, are well described by the predictions of the Standard Model. As seen, the model is built from pairs of fundamental fermions and describes their interactions as mediated by the exchange of gauge bosons. Despite the agreement between experimental results and predictions, there are many aspects that make the Standard Model not completely satisfactory:

- The large number of free parameters: they are the independent parameters of the fermion sector, three coupling constants of the gauge sectors, the two parameters describing the Higgs boson;
- The unexplained existence of 3 generations of fermions;
- The assumption of zero mass for neutrinos in contrast to experimental evidence;
- The model does not include gravitational interactions;
- The model does not provide an explanation of matter/antimatter asymmetry and Dark Matter/Energy contributions.
- The huge discrepancy between weak and gravity interaction is known as *hierarchy problem*. A fine tuning is required to deal with divergences in the Higgs sector, for scalar mass corrections evaluated at the scale where gravity interaction become relevant.

This justifies the view that the SM must be an effective low energy theory, i.e. the low energy limit of a more complete and general theory. In this approach, SM is valid up to an energy scale  $\Lambda$ , beyond which new physics appears and the inadequacy of the model is observed. The need to extend the Standard Model has given rise to numerous theories that should overcome the inconsistencies listed above.

Many models of physics beyond the Standard Model predict as possible signature for new physics the production of a heavy resonance, with the subsequent decay into a final state consisting of a pair of vector bosons ( $WW$ ,  $WZ$ ,  $ZZ$ ). These include extensions to the SM scalar sector as in the two-Higgs-doublet model (2HDM) [20] that predict new spin-0 resonances. In the Heavy Vector Triplet (HVT) phenomenological Lagrangian model [21], a new heavy vector triplet ( $W'$ ,  $Z'$ ) is introduced with the new gauge bosons degenerate in mass. Warped extra dimensions Randall–Sundrum (RS) models [22] predict spin-2 Kaluza–Klein (KK) excitations of the graviton,  $G_{KK}$  [23], and spin-0 Radions [24].

### 1.3.1 Two-Higgs Doublet Model

The Higgs mechanism led to the introduction of the Higgs boson. Since its original proposal, various extensions have been proposed that predict the existence of additional bosons besides the single one required by the minimal Higgs mechanism.

The Two-Higgs-doublet model [25, 20] introduce a second complex Higgs doublet, giving rise to five Higgs bosons:

- Two CP-even scalar fields  $h$  and  $H$ , with the latter heavier than the first one;
- One CP-odd pseudoscalar  $A$ ;
- Two charged fields  $H^\pm$ .

Two scenarios can be considered:

- The observed 125 GeV Higgs boson is the lighter scalar field  $h$ .
- The observed 125 GeV Higgs boson is the heavier scalar field  $H$ , and  $h$  has been missed.

There are four types of 2HDMs, depending on the different coupling of the two scalar fields  $h$  and  $H$  to fermions and weak gauge bosons. In type-I 2HDM all fermions couple to just one of the Higgs doublets, while in type-II the right-handed up-type fermions couple to one Higgs doublet and the right-handed down-type fermions to the other doublet. Type-III and type-IV models differ only from type-I and type-II models in their couplings to the leptons.

The Two-Higgs Doublet Model is able to explain the generation of the baryon asymmetry in the universe [26] and is also an important ingredient of axion models that can provide an explanation for the dark matter content of the universe.

### 1.3.2 Heavy Vector Triplet Model

The Heavy Vector Triplet (HVT) model [21] introduces a triplet  $\mathcal{W}$  of colorless vector bosons with zero hypercharge. It provides a broad phenomenological framework that encloses a range of different scenarios, involving new heavy gauge bosons and their couplings to Standard Model fermions and bosons.

An HVT consists of two nearly degenerate states: an electrically charged  $W'$  and a neutral  $Z'$ , collectively denoted by  $V'$ . The model allows for different coupling strengths of those states to quarks, leptons, vector bosons, and Higgs bosons by the interaction Lagrangian:

$$L_{\mathcal{W}}^{int} = -g_q W_{\mu}^a \bar{q}_k \gamma^{\mu} \frac{\sigma_a}{2} q_k - g_l W_{\mu}^a \bar{l}_k \gamma^{\mu} \frac{\sigma_a}{2} l_k - g_H (W_{\mu}^a H^{\dagger} \frac{\sigma_a}{2} i \mathcal{D}^{\mu} H + h.c) \quad (1.32)$$

where  $q_k$  and  $l_k$  are the left-handed quark and lepton doublets for fermion generation  $k$  ( $= 1, 2, 3$ );  $H$  is the Higgs doublet and  $\sigma_a$  ( $a = 1, 2, 3$ ) are the three Pauli matrices. The coupling strengths between the triplet field  $\mathcal{W}$  and the quark, lepton and Higgs fields are given by  $g_q$ ,  $g_l$  and  $g_H$  respectively. These coupling can be viewed in term of the new coupling  $g_V$ , which parametrises the interaction strength between the heavy vectors, as  $g_q = g^2 c_q / g_V$ ,  $g_f = g^2 c_f / g_V$  and  $g_H = c_H g_V$ . In the previous expressions,  $g$  is the SM  $SU(2)_L$  gauge coupling and  $c_q$ ,  $c_f$  and  $c_H$  are free parameter of the theory which can be fixed in each explicit model.

Interactions with different generations of fermions are assumed to be universal and right-handed fermions do not participate.

The  $\mathcal{W}$  field interacts with the Higgs field and thus with the longitudinally polarized  $W$  and  $Z$  bosons by means of the equivalence theorem [27, 28, 29]. The branching fractions for the decays  $W' \rightarrow WZ$ ,  $W' \rightarrow WH$ ,  $Z' \rightarrow WW$  and  $Z' \rightarrow ZH$  are therefore equal for  $V'$  masses greater than 1.5 TeV and other neutral diboson final states are either suppressed or forbidden. It allows to gain a higher sensitivity by combining not only neutral and charged diboson channels, but also eventually channels involving the Higgs boson. The decay of spin-1 boson vector into  $HH$  couple is forbidden by momentum and angular momentum conservation.

### 1.3.3 Randall–Sundrum Graviton

The Randall–Sundrum (RS) framework [22] attempts to explain the hierarchy problem by introducing *large extra dimensions* in which SM fields can propagate. This leads to Kaluza-Klein (KK) excitations of SM fields, especially including KK excitations of the gravitational field that appear as TeV-scale spin-2 Gravitons (denoted as  $G_{KK}$ ).

In some RS models the graviton has sizeable couplings to all SM fields, which do not propagate significantly into the extra dimension (bulk). This leads to large production rates in both gluon-gluon ( $gg$ ) and quark-quark ( $qq$ ) fusion modes, and substantial decay rates to diphotons and dileptons. In the “bulk RS” scenario considered in this thesis, however, the SM fields are permitted to propagate into the bulk, where they are localised. The couplings of the  $G_{KK}$  to light fermions is suppressed in this scenario, which leads to significantly reduced production rates from  $qq$  fusion and lower branching fractions to leptons and photons. The strength of the coupling depends on  $k/\overline{M}_{\text{Pl}}$ , where  $k$  is the curvature scale of the extra dimension and  $\overline{M}_{\text{Pl}} = 2.4 \times 10^8$  GeV is the effective four-dimensional Planck scale. The production cross section and decay width of the graviton scale as the square of  $k/\overline{M}_{\text{Pl}}$ . The value of the ratio and the mass of the  $G_{KK}$  are the only free parameter.

### 1.3.4 Randall–Sundrum Radion

In the RS framework the gravitational fluctuations in the usual 4-dimensional space can be viewed as the tensor fluctuations of the graviton field, while the fluctuation in the extra dimension correspond to scalar fields, known as the Radion, which are massless in the simplest scenario.

A fundamental problem in the original RS framework is that it lacks a mechanism to stabilise the radius of the compacted extra dimension,  $r_c$ . One possible mechanism to achieve this is to introduce an additional bulk scalar, which has its interactions localised on the two ends of the extra dimension [24, 30]. This causes the Radion field to acquire a mass term typically much smaller than the first KK excitation mass.

The coupling of the Radion field to SM fields scales inversely proportional to the model parameter  $\Lambda_R = \sqrt{g} k e^{-k\pi r_c} \sqrt{M_5^3/k^3}$  where  $M_5$  is the 5-dimensional Plank mass [31, 32, 33] and  $g$  the 5D metric.

The size of the extra dimension, defined as  $k\pi r_c$ , is another parameter of the model. The couplings of the Radion to fermions is proportional to the mass of the fermion while it is proportional to the square of the mass for

bosonic fields. The dominant decay mode is then into pairs of bosons when the Radion mass is above  $\sim 1$  TeV. Both the production cross-sections for the Radions and the total width scale like  $\sim 1/\Lambda_R^2$ .

### 1.3.5 Benchmark models for resonant diboson production

The benchmark models tested in this thesis work are: a simplified model predicting a heavy vector-boson triplet, a bulk RS model with a heavy spin-2 Kaluza–Klein excitation of the graviton and a bulk RS model with a heavy scalar Radion. The heavy resonance mass range investigated vary from 300 GeV up to 5 TeV.

Different resonant production mechanisms are taken into account: Gluon–gluon fusion (ggF), Drell–Yan (DY) and vector-boson fusion (VBF) processes. The Feynman diagrams of the models considered are shown in figure 1.2.

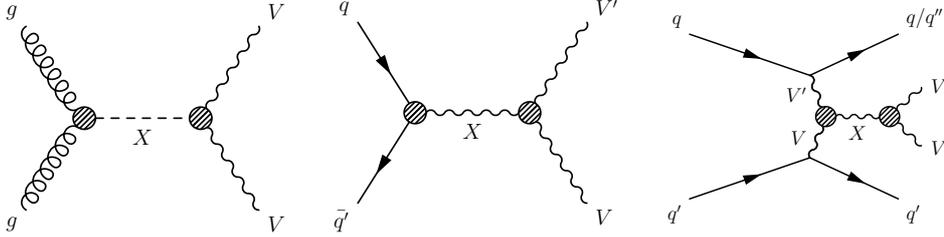


Figure 1.2: Representative Feynman diagrams for the production of heavy resonances  $X$  with their decays into a pair of vector bosons: ggF (on the left), DY (in the middle) and VBF (on the right).

Radion signals have a narrow mass width, smaller than the detector resolution (Breit-Wigner width is about 3% at 3 TeV). The widths and cross sections for Radion signals are reported in table 1.3.

A neutral heavy spin-0 Higgs boson benchmark, using the narrow-width approximation (NWA Higgs) [34], is employed in the analysis for optimization studies but not in the interpretation of the results. The characteristic of this kind of signal is the narrow width, that is dominated by the detector resolution, and the angular distributions are the same of the Radion signals. Therefore, the results using NWA Higgs signals can be compared with the

results using Radion signals. The widths and cross sections for NWA Higgs signals are reported in table 1.4.

$m_R$ (GeV)	$\Gamma$ (GeV)	$\sigma$ (fb)
800	2.1	540
1500	13	54.4
2500	57	4.3

Table 1.3: Radions cross sections in the bulk RS model via gluon–gluon fusion and the Drell–Yan process.

$m_H$ (GeV)	$\sigma_{ggF}$ (pb)	$\sigma_{VBF}$
300	6.65	1.22
500	3.08	0.47
1000	0.089	0.089

Table 1.4: Higgs production rates for ggF and VBF, calculated from Powheg-Box [35].

Three HVT scenarios are used as benchmarks for interpretation of the results. Two benchmarks are both Drell-Yan production mechanisms, while the third proceeds via the vector-boson fusion mechanism.

The first DY scenario reproduces the phenomenology of weakly coupled models based on an extended gauge symmetry, and is referred as model A [36]. The value of the coupling  $g_V$  is order one in these models.

The second DY model implements a strongly coupled scenario as in composite Higgs models, and is referred as model B [37]. This model is valid for larger values of  $g_V$ , but the width of the HVT grows with the coupling and large values of  $g_V$  do not produce a narrow resonance. Since the searches of this thesis focus on narrow resonances, the values of  $g_V$  considered are the ones in the range  $1 \leq g_V \leq 6$ , for which  $\Gamma/M$  never exceeds 10%.

In both benchmark models A and B,  $c_F$  and  $c_H$  parameters are expected to be of order one. A large coupling  $g_V$  corresponds to a small Drell-Yan production rate and, similarly, a small branching ratio into fermionic final states. Contrary, a small value of  $g_V$  leads to a small branching fraction into gauge bosons, while strongly coupled theories predict an enhanced branching ratio. Concerning the HVT coupling to SM bosons, the new heavy bosons couple dominantly to the longitudinal components of the gauge bosons and to the Higgs, while the coupling to transverse gauge bosons is generally suppressed.

Thus strongly and weakly interacting heavy vectors are expected to have a very different phenomenology: weakly coupled vectors are produced copiously, decay predominantly into two leptons or jets and have a small branching ratio into gauge bosons; strongly interacting vectors are produced less, decay predominantly into gauge bosons and two-fermion final states can be extremely rare. The  $V'$  branching fractions for decays into lepton–antilepton final states are approximately 4% in model A and only about 0.2% in model B, for each generation taken separately. The branching fractions for decays into individual diboson channels are about 2% in model A, and close to 50% in model B.

To study the rare process of vector-boson fusion a third model, model C, is designed to focus on this production mode. In this model the couplings are set to  $g_H = 1$  and  $g_F = 0$ . Model C is thus in a separate phase space domain to models A and B and assumes no Drell–Yan production. Table 1.5 reports the cross sections for the three HVT models.

$mass$ (TeV)	model A		model B		model C	
	$\sigma(W')$ (fb)	$\sigma(Z')$ (fb)	$\sigma(W')$ (fb)	$\sigma(Z')$ (fb)	$\sigma(W')$ (fb)	$\sigma(Z')$ (fb)
1.0	$2.2 \times 10^4$	$1.2 \times 10^4$	987	510	1.30	0.888
2.6	219	100	14.0	6.44	$4.78 \times 10^{-3}$	$3.14 \times 10^{-3}$
4.0	9.49	4.37	0.626	0.288	$1.27 \times 10^{-4}$	$7.92 \times 10^{-5}$

Table 1.5: Cross sections for production of  $W'$  and  $Z'$  resonances of different masses in HVT models A and B via the Drell–Yan process, in HVT model C via vector-boson fusion.

The value  $k/\overline{M}_{Pl} = 1$  is used in the RS graviton interpretations. For this choice the  $G_{KK}$  resonance width relative to its mass is approximately 6%. The decays of the gravitons in this scenario are dominated by  $G_{KK} \rightarrow t\bar{t}$ ,  $G_{KK} \rightarrow HH$ ,  $G_{KK} \rightarrow W_L W_L$  and  $G_{KK} \rightarrow Z_L Z_L$  ( $V_L$  indicates the longitudinally-polarized boson), with branching fractions that depend on mass. The  $G_{KK}$  branching fraction is largest for decays into the  $t\bar{t}$  final state, with values ranging from 42% for  $m(G_{KK}) = 0.5$  TeV to 65% for  $m(G_{KK})$  values above. Table 1.6 reports the cross sections for the bulk RS graviton model.

Figure 1.3 shows the diboson invariant mass and the polar angle of hadronically-decaying  $V$  boson in the rest frame for different signal hypotheses.

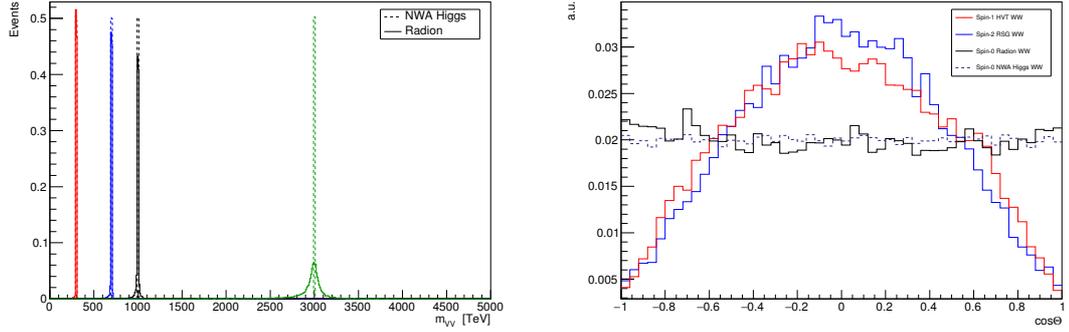


Figure 1.3: Diboson invariant mass for Radion and heavy Higgs signals, for different mass hypotheses (on the left);  $\cos\Theta$  for different signal hypotheses, where  $\Theta$  is the the polar angle of hadronically-decaying V boson in the rest frame of the hypothetical new particle X (on the right).

$mass$ (TeV)	$\sigma(G_{KK})$ (fb)
1.0	5.83
2.6	1.41
4.0	$3.25 \times 10^{-2}$

Table 1.6: Gravitons cross sections in the bulk RS model via gluon–gluon fusion and the Drell–Yan process.

## Chapter 2

# The ATLAS experiment at the LHC

The *Large Hadron Collider* (LHC) [38, 39] is a particles accelerator in which two protons beams collide at very high energy. The particles produced in these collisions are detected by four big experiments:

- ATLAS [40] (**A Toroidal LHC Apparatus**);
- CMS [41] (**C**ompact **M**uon **S**olenoid);
- LHCb [42];
- ALICE [43] (**A Large Ion Collider Experiment**).

ATLAS and CMS have been designed in order to investigate many Particle Physics arguments, focusing on the Standard Model, the study of the properties of the Higgs boson and searches of evidence of new Physics Beyond the Standard Model.

LHCb studies the b-quark physics and investigate the outstanding problem of the matter-antimatter asymmetry in the universe.

ALICE focuses on the quark gluon plasma searches, by heavy ion collisions studies.

### 2.1 The LHC storage ring

The Large Hadron Collider is a proton-proton collider located inside the same tunnel used also for the LEP accelerator [44] at the CERN of Geneva (fig. 2.1).

The maximum energy that each of the protons beams can reach for frontal

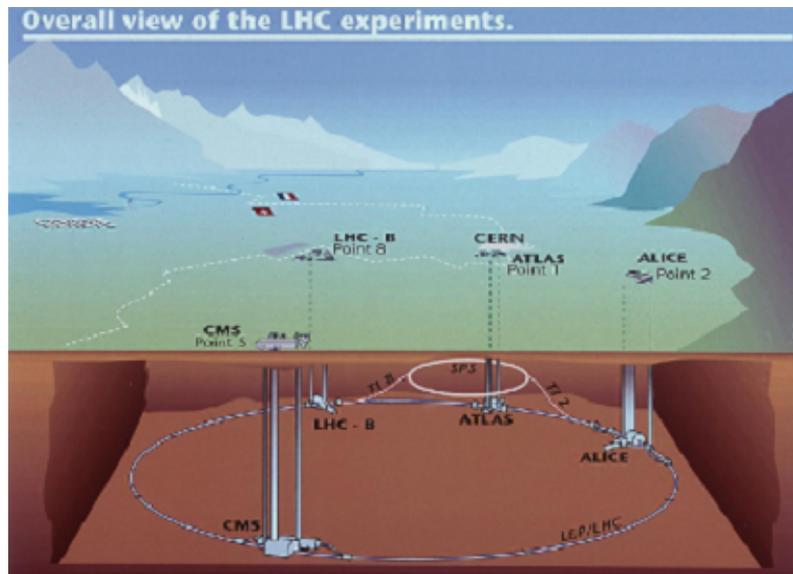


Figure 2.1: Schematic view of the underground zone in which LHC is situated.

collisions (head-on) is 7 TeV, with a project luminosity of  $L = 10^{35} \text{cm}^{-2} \text{s}^{-1}$ . LHC is made by two superconductor rings, with a circumference of 26.7 km, used to accelerate and to collide hadrons. The whole LHC experimental area extends on the French-Swiss border near Geneva, where CERN is based. In order to restrict the costs, some of the existing infrastructure before LHC have been reused where it was possible.

### 2.1.1 Technical specifications

#### Structure

LHC is a particle-particle collider, so unlike the particle-antiparticle collider, it's composed by two rings. The beams interact in four points where the particle detectors are installed. The main accelerator components are the magnetic dipoles and quadrupoles and the resonant cavities. The cryogenic super-conductive dipoles work at the temperature of 1.9 K and they have been build in order to produce magnetic fields of 8.33 T, so that protons can travel on circular orbits. To cool them, super-fluid helium is used. The quadrupoles are used in order to focus the beam, while the resonant cavities accelerate particles. The two LHC rings are divided in eight octants made by arches and straight sections of approximately 528 meters. The two zones with high luminosity are situated in two diametrically opposed straight sections.

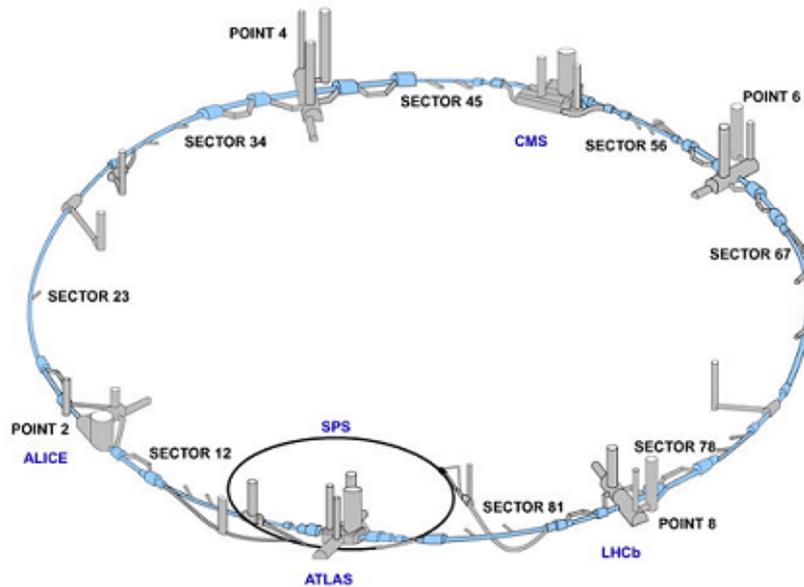


Figure 2.2: Experiments arrangement in LHC.

Here the ATLAS and CMS experiments are situated: Point 1 and Point 5 (fig. 2.2). The other two experiments, ALICE and LHCb, are situated in the two zones where LHC reach the minimum luminosity: Point 2 and Point 8 (fig. 2.2). In the remaining four straight sections there aren't more beams intersections. Injections zones are in the Point 2 and 8, for the injection of protons bunches in clockwise and counter-clockwise respectively. Points 3 and 7 contain collimation systems while Point 4 contains two independent radio frequency (RF) systems, one for each beam. The straight section at Point 6 contains the beam dump insertion: this operation is performed using a combination of horizontally deflecting fast-pulsed ('kicker') magnets and vertically-deflecting double steel septum magnets.

### Acceleration mechanism

Beams acceleration process is divided in stages, according to the different type of devices that the beam crosses. Injectors chain is showed in figure 2.3. From the protons source to the last acceleration step we have:

- LINAC2;
- Proton Synchrotron Booster (PSB);

- Proton Synchrotron (PS);
- Super Proton Synchrotron (SPS);
- LHC.

A protons beam is produced from  $H_2$  gas and protons are then accelerated at 300 mA beam current. Protons enter in the PSB with an energy of 50 eV from LINAC2 accelerator, and here they are accelerated up to 1.4 GeV. PS accelerates protons up to 25 GeV and then the beams are injected in the SPS, where the protons reach an energy of 450 GeV.

As last step, the two beams enter in LHC where they circulate in opposite directions until they reach the energy required for the collisions. Beams are accelerated thanks to eight resonant cavities, whose electric fields oscillate at 400.8 MHz, in order to kick protons and compensate for losses.

Protons enter in bunches at LHC and, to avoid colliding with gas molecules inside the accelerator, ultra-high vacuum is needed for the pipes in which particle bunches travel. Each beam is composed by 2808 bunches with a 25ns separation and  $\sim 10^{11}$  protons per bunch.

At LHC the *pile-up* plays an important role. When a collision between protons bunches happens, a potential interesting event is covered by other non-interesting events produced in the same collision. Two kind of pile-up exists: pile-up *in-time* due to additional proton-proton interactions in the same bunch-crossing, and pile-up *out-of-time* due to energy deposits in the calorimeter due to following bunch-crossing.

### Machine luminosity

The number of events per second produced in the collisions is:

$$N_{ev} = L \cdot \sigma_{ev} \quad (2.1)$$

where  $\sigma_{ev}$  is the cross section of the event and  $L$  is the machine luminosity, that depends only on the beam parameters:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (2.2)$$

where  $N_b$  and  $n_b$  are respectively the number of particles per bunch and the number of bunches per beam,  $f_{rev}$  is the revolution frequency,  $\gamma_r$  is the gamma relativistic factor;  $\epsilon_n$  is defined as the product of the widths of the positions distribution of the particles in the bunch ( $\sigma_x$ ) and the momentum distribution of the particles in the bunch ( $\sigma_p$ ), while  $\beta^* = \frac{\sigma_x}{\sigma_p}$  is the ratio of

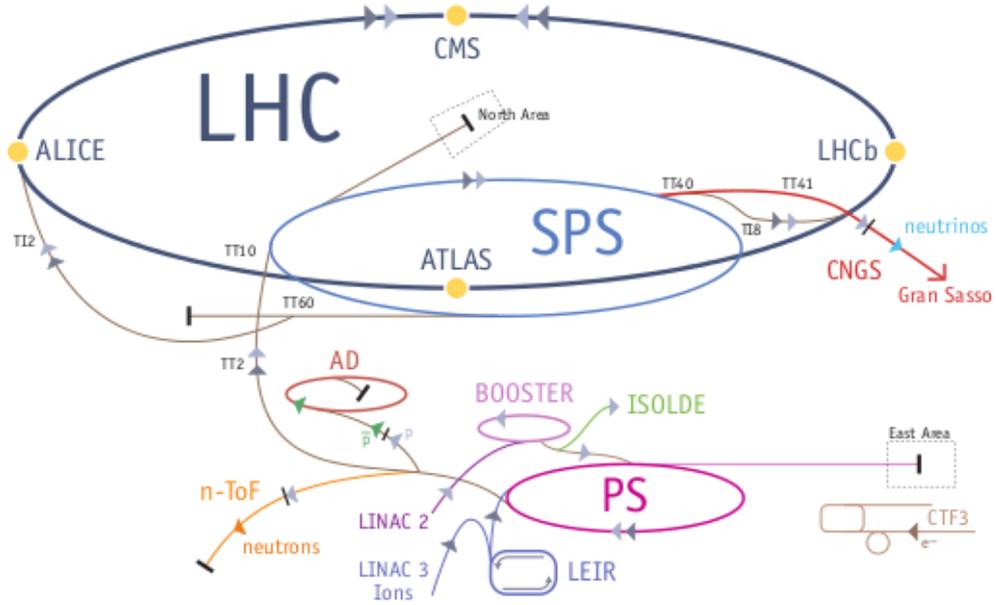


Figure 2.3: LHC acceleration and pre-acceleration devices schema.

these widths. Finally  $F$  is the geometric luminosity reduction factor due to the crossing angle at the interaction point.

LHC luminosity is not constant on the whole data taking cycle, but it decrease due to a degradation of the intensity of the beam. The main cause of the luminosity decrease is the beam lost due to the collisions in the interaction points. The initial decay constant that arise from this effect is:

$$\tau_{nuclear} = \frac{N_{tot}(0)}{L\sigma_{tot}k} \quad (2.3)$$

where  $N_{tot}(0)$  is the initial beam intensity,  $L$  the initial luminosity,  $\sigma_{tot}$  the total cross section ( $\sigma_{tot} \simeq 110$  mb at 13 TeV) and  $k$  is the number of interaction points.

Other contributions come from particles loss due to radiations: scattering of the particles with the gas inside the experiment detectors, non-linear forces during the beam interactions, intra-beam scattering (IBS).

If the luminosity decrease is approximated by an exponential function, the decay constant can be written as:

$$\frac{1}{\tau_L} = \frac{1}{\tau_{IBS}} + \frac{1}{\tau_{gas}} + \frac{1}{\tau_{nuclear}} \quad (2.4)$$

With the following assumptions:

$$\tau_{IBS} = 80 \text{ h}$$

$$\tau_{gas} = 100 \text{ h}$$

$$\tau_{nuclear} = 29 \text{ h}$$

the resulting decay constant is

$$\tau_L = 14.9 \text{ h} \quad (2.5)$$

Integrating the luminosity over a run:

$$L_{int} = L_0 \tau_L [1 - e^{-T_{run}/\tau_L}] \quad (2.6)$$

where  $T_{run}$  is the data taking period.

## 2.1.2 Data-taking

### Run I (2009 – 2012)

First LHC operations started in the September 2008 but they were interrupted since a mechanical damage occurs in the dipole and quadrupole system. The fix of this damage took more than one year, and in November 2009 LHC resumed the operations starting with proton-proton collisions at 900 GeV centre of mass energy. From March 2010 to the end of 2011 the collisions have been performed at  $\sqrt{s} = 7$  TeV and from the 2012 the energy has been increased up to 8 TeV. At the end of the 2012 and until the 2015 LHC was turned off and upgrades on the apparatus have been performed. Figure 2.4 reports the integrated luminosity delivered by LHC during the 2011-2012 data-taking period.

The 4th July of 2012 ATLAS and CMS collaborations have announced the discovery of a Higgs-like particle, the last fundamental particle foreseen in the SM.

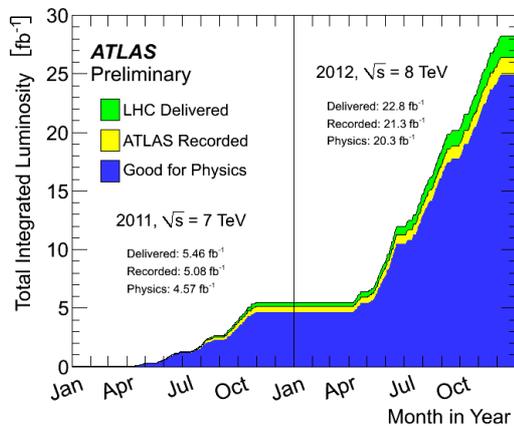


Figure 2.4: Integrated luminosity during stable beams for pp collisions in the *Run I* versus time delivered by LHC (green), recorded by ATLAS (yellow), and certified to be good data for physics (blue).

### Run II (2015 – 2018)

LHC operations restarted in April 2015 at  $\sqrt{s} = 13$  TeV. This second phase is called *Run II* [45] and covered the period 2015-2018.

The data-taking conditions improved significantly respect to the *Run I*: the peak luminosity increased from  $5$  to  $19 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$ , as the number of bunches and average bunch current were increased and also the transverse beam sizes near the interaction points have been reduced.

The total luminosity recorded during the *Run II* is around  $140 \text{fb}^{-1}$  and figure 2.5 reports the integrated luminosity delivered by LHC, while figure 2.6 reports the comparisons of *Run I* and *Run II* integrated luminosities.

## 2.2 The ATLAS experiment

ATLAS is a multi-purpose experiment with searches varying from Standard Model precision measurements to search for evidence of new physics. ATLAS is a big collaboration of about 3000 scientific authors from 183 institutions around the world. The experimental apparatus has the following structure:

- An Inner Detector (ID), covered by a solenoidal superconducting magnet, that measures the direction, momentum and charge of interacting particles.
- Electromagnetic and hadronic calorimeters that measure the energy deposited by the particles and for the jets reconstruction.

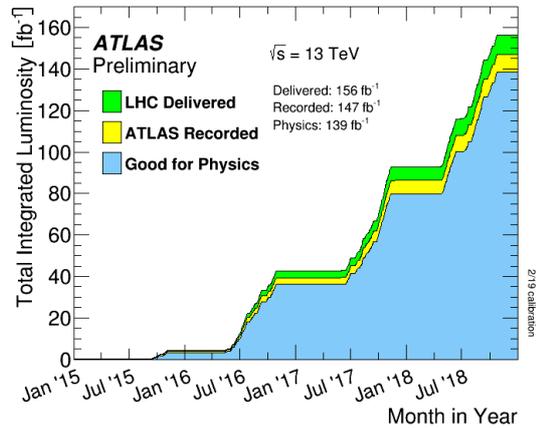


Figure 2.5: Integrated luminosity during stable beams for pp collisions in the *Run II* versus time delivered by LHC (green), recorded by ATLAS (yellow), and certified to be good data for physic (blue).

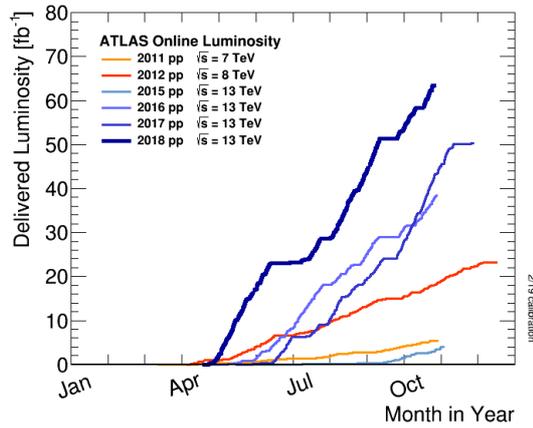


Figure 2.6: Comparisons of *Run I* and *Run II* integrated luminosities versus day delivered by LHC.

- Muon Spectrometer (MS) that identifies and reconstructs muons. A toroidal magnet system is used in order to provide the magnetic field to the MS.
- A trigger system, to reduce data-acquisition rate and select only potentially interesting physics events.
- A data acquisition system, to store the interesting events.

The detectors are required to have peculiar properties in order to cover the wide range of physics phenomena studied with the ATLAS detector:

- The detector needs a fast electronic due to the special LHC conditions;
- High granularity is required to reconstruct properly the events and avoid overlaps;
- A large acceptance in pseudorapidity ( $\eta$ ) and an almost full azimuthal angle coverage is required;
- The inner detectors are required to have a good resolution in the momentum and secondary vertex reconstruction;
- A very good calorimetric system is needed in order to identify and measure the energies of electrons, photons, jets and missing energy;
- An optimal identification and the momentum resolution of the muons is required;
- The trigger system has to be fast and highly efficient in order to maximize the background events rejection.

Figure 2.7 shows the full detector: ATLAS has the dimensions of a cylinder, it's 46 m long, 25 m tall and weights 7000 tons.

The tightness of the experimental apparatus allows to detect the production of neutrinos or other unknown particles non directly detected in the apparatus, by reconstructing missing energy and momentum.

In the following sections will be described with more details the different detectors that constitute the ATLAS experiment.

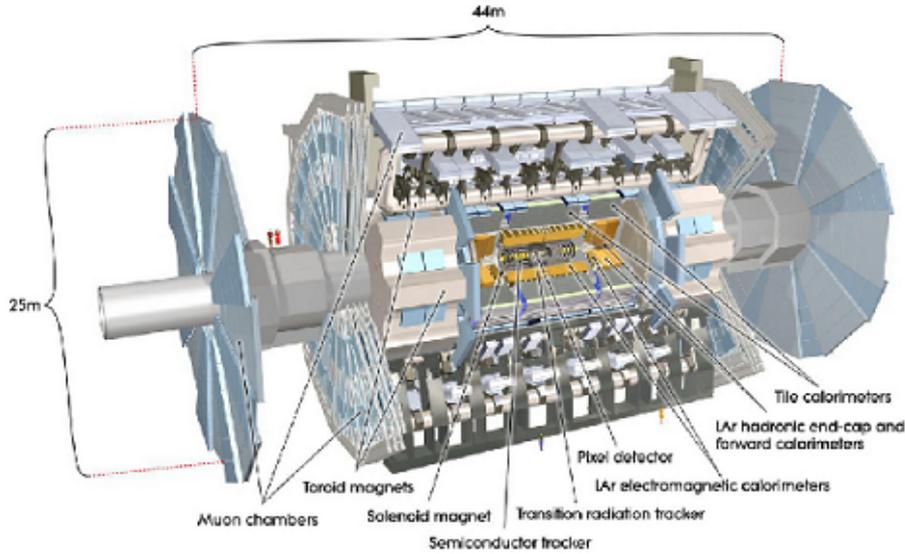


Figure 2.7: The ATLAS apparatus at LHC.

### 2.2.1 ATLAS Coordinate System

Two coordinate systems are used in the ATLAS experiment: a cartesian right-handed coordinate system for geometry studies and a spherical coordinate system, more useful in the physics analyses. The origin of the coordinate system is the nominal beams interaction point, with the  $z$  axis along the beam direction and the  $x - y$  plane that is transverse to the beam direction. The  $x$  axis points to the centre of the LHC ring while the  $y$  axis goes upwards. The interaction point defines two regions, the one with  $z > 0$  is called *side A* and the one with  $z < 0$  is called *side C*. The  $z = 0$  plane is the *side B* (figure 2.8).

In the spherical coordinate system, the azimuthal angle  $\phi \in [-\pi, \pi]$  is measured in the  $x - y$  plane and it's defined as:

$$\phi = \frac{1}{\tan(x/y)} \quad (2.7)$$

The azimuthal angle  $\phi$  is equal to zero along the  $x$  axis and increases in clockwise direction in positive  $z$  direction. The polar angle  $\theta$  is defined in the  $y - z$  plane and it's equal to zero along the  $z$  axis.

Since the polar angle is not invariant under Lorentz boost along the  $z$  direction, another variable called *pseudorapidity* is used (fig 2.9):

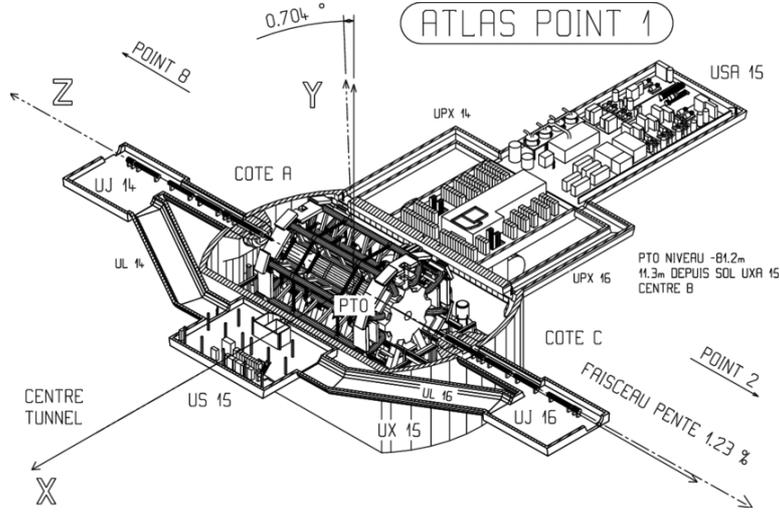


Figure 2.8: ATLAS coordinate system.

$$\eta = -\ln \left[ \tan\left(\frac{\theta}{2}\right) \right] \quad (2.8)$$

The detector is divided in two regions:

- Barrel region (BR)  $|\eta| < 1.05$ , that corresponds to the central detector region, so more close to the interaction point;
- Endcap region (ER)  $1.05 < |\eta| < 2.7$ , that includes the left and right forward apparatus.

For massive objects, like jet, *rapidity* is used:

$$y = \frac{1}{2} \ln \left[ \frac{E + p_L}{E - p_L} \right] \quad (2.9)$$

where  $p_L$  is the longitudinal momentum component.

The angular separation in the  $\eta - \phi$  plane, defined as:

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.10)$$

is used to measure the angular distance between two tracks.

LHC is an hadronic collider, so non-elementary particles made of partons collide. The effective centre of mass interaction energy depends by the

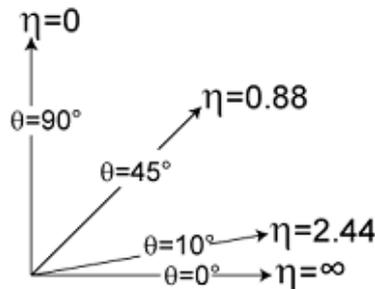


Figure 2.9: Pseudorapidity  $\eta$  for different  $\theta$  values.

momentum fraction carried by the partons that actually participate to the hard scattering process, so it is not exactly known. It is natural to study the kinematics of the interaction in the transverse  $x - y$  plane, in which the energy conservation is still valid (since the transverse momentum component of the partons is negligible respect to the longitudinal one).

## 2.3 ATLAS experimental apparatus

### 2.3.1 Magnetic system

The main feature of the ATLAS detector is the magnetic system. It's composed by 4 superconducting magnets, it's 26 m long and has a diameter of 20 m. It's composed by (fig. 2.10):

- A central solenoid (CS), symmetric along the beams direction, that covers the central region of the detector and provides an uniform magnetic field of 2 T along the  $z$  axis. The CS is 5.3 m long and with a diameter of 2.4 m. The solenoidal field bends tracks of the particles of the Inner Detector in the transverse plane in order to measure the particle transverse momentum.
- A system of 3 superconducting toroids (one in the barrel and two in the end-caps) that surround the CS and provides a magnetic field, whose peak intensities are 3.9 T and 4.1 T in the central and forward region respectively. This system is one of the peculiarities of the ATLAS detector and its aim is to have a large spatial acceptance for the measurement of muons transverse momentum. Each toroid consists of 8 rectangular coils radially and symmetrically assembled around the beam axis. The end-cap toroid system is rotated by  $22.5^\circ$  with respect to the barrel system in order to have a radial overlap.

The double magnetic system provides two independent measurements of the muon momentum in the inner detector and in the muon spectrometer, and this ensures good muon momentum resolutions from few GeV up to the TeV scale.

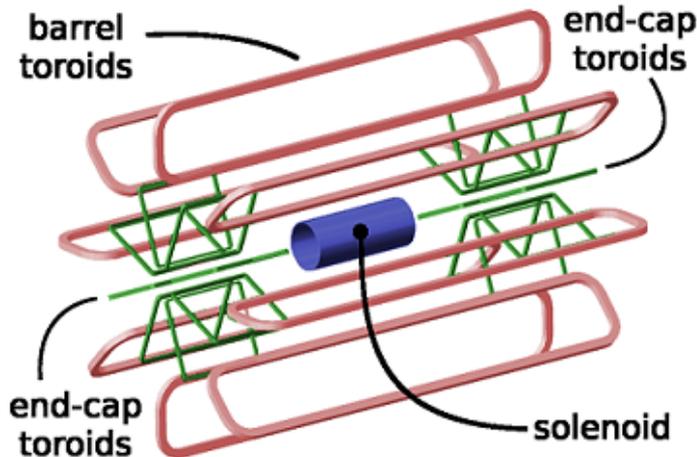


Figure 2.10: ATLAS magnetic system schema.

### 2.3.2 Inner detector

The Inner Detector (ID) [46] has been designed in order to have a good pattern recognition, an excellent momentum resolution and a good primary and secondary vertex reconstruction for charged tracks with transverse momentum of more than 200 MeV.

To meet these requests, detectors with high granularity are required (due to the enormous tracks density expected at LHC). The ID is composed by three detectors (fig. 2.11) surrounded in a magnetic field parallel to the beam axis: *Pixel Detector*, *Semiconductor Tracker (SCT)* and *Transition Radiation Tracker (TRT)*. In the central part, three layers of high-resolution pixel detectors and eight semiconductor trackers are installed. The total number of layers has to be limited due to the high amount of material (in order to avoid multiple scattering processes) and to keep costs down. The combination with the Transition Radiation Tracker, installed in the outer part of the ID, allows to obtain an high resolution in both the  $z$  and  $\phi$  directions. The Inner Detector covers the  $|\eta| < 2.5$  region.

The external radius of the tracking cavity is 155 cm long and the total length is 7 m. From a mechanical point of view, the ID is composed by 3

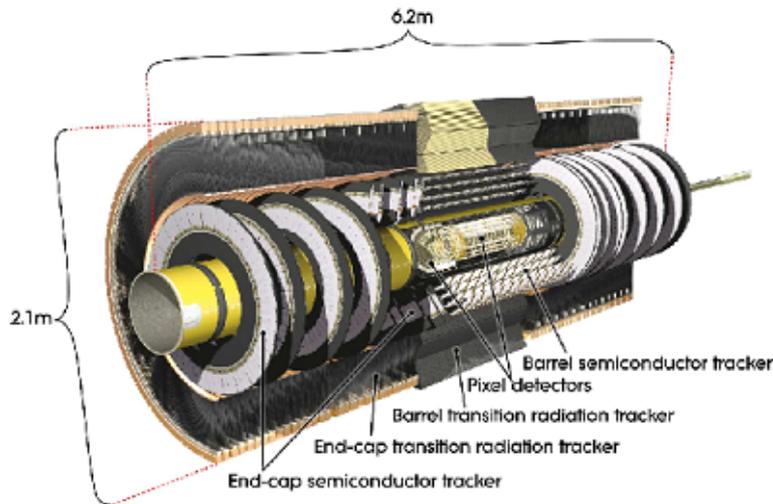


Figure 2.11: The ATLAS Inner Detector.

units: a cylindrical part that covers  $\pm 80$  cm from the interaction point, and two identical end-cap sides that cover the remaining cylindrical cavity. In the central part, the high-precision detector layers are installed in a concentric cylinders around the beam axis, while in the detectors in the end-cap regions are installed in cylinders orthogonal to the beam axis.

The Inner Detector is the part of the ATLAS detector more close to the interaction point, so more prone to radiation damages, and thus a frequent maintenance is necessary in order to guarantee high performance.

### Pixel detector

The pixel detector allows to perform high-precision measurements, to determine with a very good resolution the impact parameter and to reconstruct the tracks of the particles with short lifetimes, like  $\tau$  leptons and  $b$ -quarks. The pixels are installed on concentric cylinders around the beam axis and on cylinders orthogonal to the beam axis in the end-cap regions. The pixels are arranged so that the tracks from the interaction point cross at least 3 pixel layers. The system consists of 3 cylindrical bodies with an average radius of about 4, 11 and 14 cm, and 5 discs per side with a radius between 11 and 20 cm, which complete the angular coverage. 140 millions of pixels detectors, with a  $50 \mu\text{m}$  step in the  $R - \phi$  direction and  $300 \mu\text{m}$  in the  $z$  direction are installed. The spatial resolutions are (averaged on pseudorapidity):

$$\sigma(R - \phi) \simeq 12 \mu\text{m} \quad \text{for all pixels}$$

$$\sigma(z) \simeq 66\mu\text{m} \quad \text{in the barrel region}$$

$$\sigma(R) \simeq 77\mu\text{m} \quad \text{for the discs}$$

Electronic modules for the detectors readout have big areas, and each pixel module has a single circuit. Moreover, they are equipped with pattern buffers, in order to store the data while waiting for the level-1 trigger decision.

### Semiconductor Tracker

The Semiconductor Tracker (SCT) is installed in the radial mid-zone (fig. 2.12); it provides at least 4 precision measurements for track and contributes to the momentum, impact parameter and vertex position measurements, as well as good pattern recognition, thanks to the detector high-granularity. 8 silicon micro-strip detector layers provide in the barrel region the  $R - \phi$  and  $z$  coordinates. The sizes of each device are of  $6.36 \times 6.40 \text{ cm}^2$  with 768 readout strips in steps of  $80 \mu\text{m}$ . The whole detector includes  $61 \text{ m}^2$  of silicon detectors and a total of 6.2 millions of readout channels. The spatial resolution is  $16 \mu\text{m}$  in  $R - \phi$  and  $580 \mu\text{m}$  in  $z$ .

Due to the high radioactivity, the SCT needs to work at low temperature (between the  $-5$  and  $-10 \text{ }^\circ\text{C}$ ) and so has been constructed with materials with a low thermal expansion coefficient and integrated with a robust cooling system.

### Transition radiation tracker

The TRT provides a large number of additional hits, around 36 per track. The detector is made of straw tubes with 4 mm diameter, able to operate at high particle fluxes thanks to the detecting wires, well isolated within single gas volumes. The detector tubes are surrounded by a propylene foam. The TRT works with an not inflammable gas mixture composed of 70%  $Xe$ , 20%  $CO_2$  and 10%  $CF_4$ .

When a charged particle crosses the detector, it generates a transition radiation due to the different dielectric constants of air and propylene. The best electrons collection time is  $\sim 48 \text{ ns}$ , and the spatial resolution coming from the drift time is  $\sim 130 \mu\text{m}$ . TRT position is showed in figure 2.12.

### 2.3.3 Calorimetric system

The calorimeters (fig. 2.13) are placed between the Inner Detector and the Muon Spectrometer. They are sampling calorimeters composed of layers of

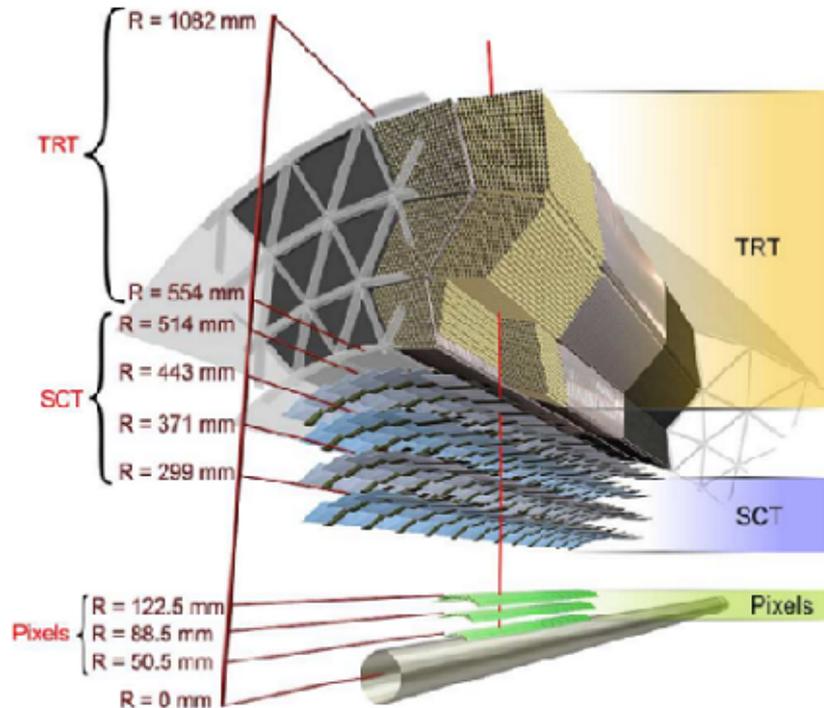


Figure 2.12: 3D schema of the 3 ID tracks detectors.

active material and passive (or absorber) material. A particle that crosses the passive layers interacts with the material, and produces a shower of secondary particles that are detected by the active material. The particles of the shower gradually lose their energies crossing the calorimeter, and finally they are completely absorbed.

Two different types of calorimeters are employed, in order to give a good resolution and reconstruction of the electromagnetic showers and to contain the hadronic showers, avoiding their passage into the Muon Spectrometer.

The whole calorimetric system [47] is composed of:

- An electromagnetic calorimeter (EM), that covers the  $|\eta| < 3.2$  region;
- A cylindrical hadronic calorimeter (HC *Tile Barrel*), that covers the  $|\eta| < 1.7$  region;
- Two end-cap hadronic calorimeters (HEC), that cover the  $1.5 < |\eta| < 3.2$  region;

- Two forward calorimeters (FCAL), that cover the  $3.2 < |\eta| < 4.9$  region.

The electromagnetic calorimeter is a lead and liquid argon (LAr) detector with an accordion geometry. A pre-sampler detector, that corrects the measures for the energy lost in the material, is installed just behind the cryostat cold wall.

The active material of the hadronic calorimeter (*Tile Barrel*) consists of plastic scintillator plates surrounded in an iron absorber. It's divided in 3 sections: a central cylinder and two extended cylinders.

The end-cap calorimeters are made of copper (Cu) and liquid argon (LAr), while the FCAL are made of liquid argon and rod-shaped electrodes in a tungsten matrix.

When an high-energy particle, like an electron or a photon, interacts with the material, it generates an electromagnetic shower composed by secondary electrons and photons. The initial energy of the interacting particle is divided between all the secondary particles and deposited in the material in form of ionization. By detecting and measuring the energy deposited in the material, it's possible to determine the kind of the interacting particle and to measure its energy.

Particles that are subject to the strong interaction, impacting on a material, interact with the nuclei of the material and produce an hadronic shower. As in the electromagnetic case, measuring the energy of the secondary particles it's possible to determine the energy of the starting interacting particle.

The calorimeters are build so that the electrons and the photons are completely absorbed in the electromagnetic calorimeter, while an hadronic particle crosses quite undisturbed the electromagnetic calorimeter and then it's absorbed in the hadronic calorimeter.

### Electromagnetic calorimeter (ECAL)

The magnitude that characterizes an electromagnetic calorimeter is the material radiation length  $X_0$ ; it's defined as the average distance that reduces the interacting particle energy of a factor  $1/e$ .

The physic processes that lead to a production of an electromagnetic shower are the  $e^+e^-$  couples creation, starting from a photon, and a photon emission by an electron (*bremstrahlung*). These processes can happen only through the interaction with a second body, like an atomic nuclei (fig. 2.14).

The combination of couple creation and bremstrahlung leads to the development of a shower made by positrons, electrons and photons (fig. 2.15).

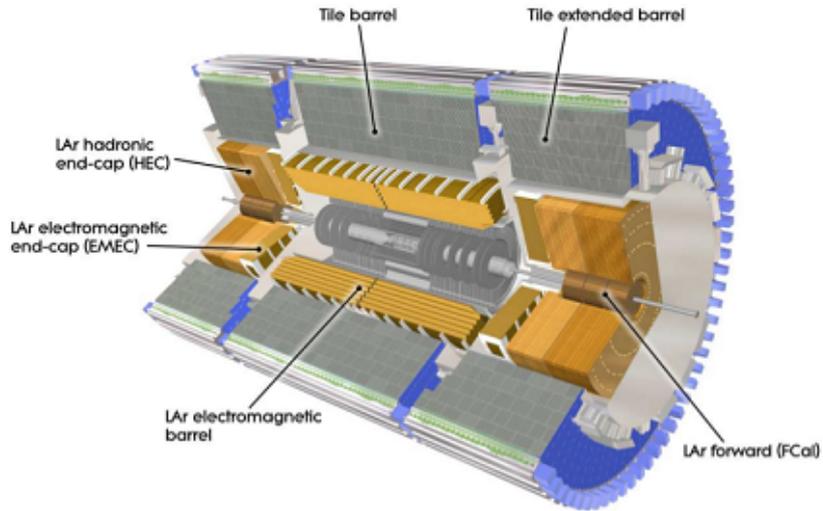


Figure 2.13: ATLAS calorimetric system schema.



Figure 2.14: Couple creation and bremsstrahlung through the interaction with an atomic nuclei.

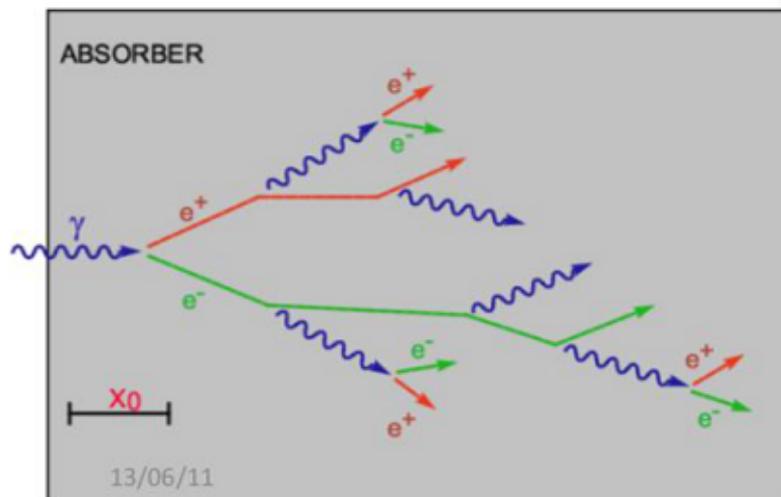


Figure 2.15: Development of an electromagnetic shower, radiation length  $X_0$  is showed.

The concatenation of couple creation and bremsstrahlung processes continues until the energy for the couple creation doesn't go below a certain threshold. At this point the  $e^+e^-$  couple loses energy preferentially by atomic collisions, rather than bremsstrahlung emissions, arresting so the shower. Crossing a layer  $X_0$  thick, an electron loses on average  $2/3$  of its energy radiating a photon, with a probability of  $7/9$  to generate an  $e^+e^-$  couple. So, as first approximation, it's possible to assume that every  $X_0$  a new generation of particles in the shower is produced.

The structure of the electromagnetic calorimeter system is showed in figure 2.16; it is designed in order to maximize the energy resolution. In the central region ( $|\eta| < 2.5$ ), the EM calorimeter is composed by 3 different layers: the inner one is very small in length,  $4.3 X_0$  thickness, but with very high granularity in  $\eta$  and  $\phi$  ( $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ ), so that it's able to discriminate the  $\gamma$  from the  $\pi^0$ . The second region, of  $16 X_0$  thickness, is segmented in cells with dimensions  $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ , that allows to have a good spatial resolution for the reconstructed shower. The third layer,  $2 X_0$  thickness, has a segmentation of  $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$  and it's dedicated to high energy electrons and photons.

The longitudinal sampling of the showers is achieved by replicating 4 times the cell structure in the radial direction. The total length is around  $25 X_0$  in the barrel region and around  $26 X_0$  in the end-cap regions with 190000 readout channels in total.

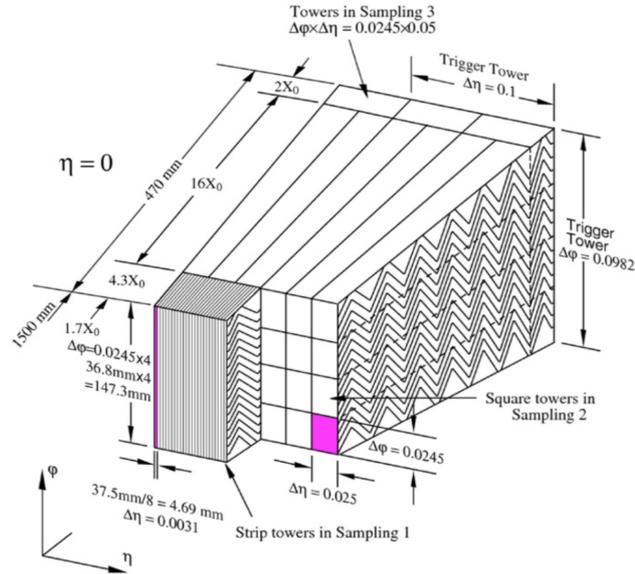


Figure 2.16: Electromagnetic calorimeter system.

### Hadronic calorimeter (HCAL)

ATLAS hadronic calorimeters cover the  $|\eta| < 4.9$  region, adopting different techniques for different pseudorapidity ranges. The thickness is an important feature to be taken into account for the construction of the hadronic calorimeter, since it has to be able to minimize the particles that reach the muon spectrometer, and it has to provide a good resolution for high-energy jets and a good measure of the missing transverse energy. The interaction length,  $\lambda_I$ , is defined in the hadronic calorimeter in the same way of the radiation length  $X_0$  for the electromagnetic calorimeter. The ATLAS hadronic calorimeter is composed by:

**Tile Calorimeter:** It's a sampling calorimeter where the iron is used as absorber material and the scintillator as active material. It covers the  $|\eta| < 1.7$  region and the thickness is about  $10 \lambda_I$  for  $|\eta| = 0$  with a granularity of  $\Delta\eta \times \Delta\phi \simeq 0.1 \times 0.1$ .

**Hadronic end-cap calorimeter (HEC):** In this calorimeter copper plates are used as absorber and LAr as active material. It covers the  $1.5 < |\eta| < 3.2$  region, with an overlap with the *Tile calorimeter* and the *Forward calorimeter*. The granularity is  $\Delta\eta \times \Delta\phi \simeq 0.1 \times 0.1$ .

**Forward calorimeter (FCAL):** This calorimeter has a complex de-

sign due to the high radiation flux that is present in the region where is placed (around 4.7 m from the interaction point). It covers the  $3.1 < |\eta| < 4.9$  region and it's able to reduce background particles of the Muon Spectrometer. It consists of 3 modules for side, all of them with LAr as active material: FCAL1, FCAL2, FCAL3. The first module is optimized for electromagnetic measures, while the other modules measure mainly the energies of the hadronic interactions using tungsten as absorber.

### Performance of the calorimeters

The energy resolution of the electromagnetic calorimeter comes from the sum in quadrature of independent terms:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (2.11)$$

where  $a$  is the sampling factor (that includes also the statistical fluctuations),  $b$  is the factor that accounts for the noise coming from the electronic and the signal pile-up, and  $c$  is a constant that accounts for mechanical effects, calibrations and non-uniform sources which arise in systematic errors. The energy resolution of the ATLAS EM calorimeter depends of the values of these parameters, and the expected resolution between 2 GeV and 5 GeV is:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.3\%. \quad (2.12)$$

The expected angular resolution is  $40 \text{ mrad}/\sqrt{E[\text{GeV}]}$ , allowing a good measurement of the shower direction along  $\eta$ .

For the hadronic calorimeters the energy resolution can be expressed as:

$$\frac{\sigma_E}{E} = \sqrt{\frac{c_{int}^2 + c_{samp}^2}{E}} \oplus a \quad (2.13)$$

where  $a$  describes the non-Gaussian fluctuations of electromagnetic showers,  $c_{int}$  represents Gaussian fluctuation of the initial energy and  $c_{samp}$  accounts for statistical and sampling fluctuations.

The expected resolutions are:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \quad \text{for } |\eta| < 3.0; \quad (2.14)$$

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \quad \text{for } 3.0 < |\eta| < 4.9. \quad (2.15)$$

### 2.3.4 Muon Spectrometer

Unlike electrons, positrons and hadrons that lose all their energy in the calorimeters (ECAL and HCAL), muons are *minimum ionizing particles* (MIP) and they cross all the detector, losing only a small fraction of their energy. The Muon Spectrometer [48] constitutes the outer part of the ATLAS detector and allows to identify muons from 3 GeV (at lowest energies the particle is absorbed by the calorimeter system) up to 1 TeV. The Muon Spectrometer is based on the deflection of the muons trajectories by the magnetic toroidal system and it's divided in the barrel region (that covers  $|\eta| < 1.4$ ) and in the end-caps regions ( $1.6 < |\eta| < 2.7$ ). In the transition region  $1.4 < |\eta| < 1.6$  the deflection is produced by the combination of the two magnetic fields in the barrel and end-cap regions.

This configuration allows to have a field mainly orthogonal to the muon trajectories, in order to measure the transverse momentum ( $p_T$ ). The spectrometer consists of:

- High precision tracking chambers, composed by Monitored Drift Tubes (MDT) in the barrel region and Cathode Strip Chambers (CSC) in the end-cap regions;
- Trigger system, that provides informations on the tracks that cross the detector, employing a fast electronic and composed by Resistive Plate Chamber (RPC) in the barrel region and Thin Gap Chambers (TGC) in the end-caps.

A schema of the muon spectrometer is showed in figure 2.17.

In the barrel region the tracks are measured in chambers placed in 3 cylindrical layers around the beam axis; in the transition and in the end-cap regions the chambers are installed in 3 layers placed orthogonally the beam axis.

The *Monitored Drift Tubes* (MDT) provides precision measurements of the tracks momentum on almost the whole pseudorapidity range. They are chambers made of 3-4 layers of drift tubes with 15 mm radius, filled with a gas mixture  $Ar - CO_2$  that is ionized with the muon passage. Each tube provides a single measure (hit) and the tube multiplicity improves the track reconstruction. Inside the tube is placed an anode wire of 50  $\mu\text{m}$  diameter.

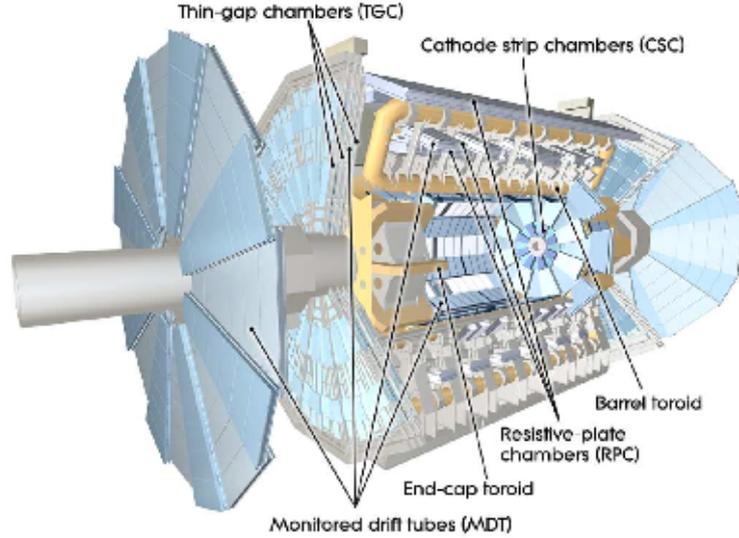


Figure 2.17: Layout of the muon spectrometer and disposition of the different chambers inside the spectrometer.

A charged particle that crosses the tube loses energy by ionization, and produces electron-positron couples. The electrons go towards the anode wire and an avalanche multiplication process takes place near the wire. The time of arrival is recorded, so the distance of the muon track from the wire can be obtained providing a spatial resolution of  $\sim 80 \mu\text{m}$ . The collected charge is proportional to the initial particle energy. The chamber momentum resolution is less than 10% for muons with a momentum between 10 GeV and 1 TeV.

In the  $2 < |\eta| < 2.7$  region, the MDT are made of *Cathode Strip Chambers* (CSC). They are multi-wire proportional chambers with the cathode segmented in strips, more resistant to the high radiation flux intensity. Track position is obtained by measuring the charge induced on the cathode, due to the avalanche generated on the anode. The drift time is less than 25 ns, while spatial resolution is  $\sim 50 \mu\text{m}$ .

The trigger system covers the  $|\eta| < 2.4$  region. An RPC chamber [49] is composed by a mixture gas of tetrafluoroethane ( $C_2H_2F_4$ ), isobutane ( $(CH_3)_3CH$ ) and sulphur hexafluoride ( $SF_6$ ), with a percentage of 94.7%, 5.0% and 0.3% respectively, contained between two parallel Bakelite plates 2 mm spaced by an insulating spacer under an electric field of typically 4.5

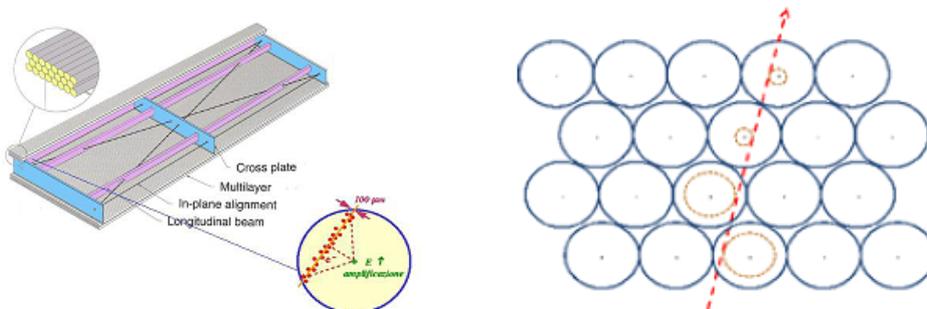


Figure 2.18: MDT chamber structure (on the left) and schema of the muon trajectory measurement with the MDT by drift time (on the right).

kV/mm. When a particle crosses an RPC chamber, the primary ionizing electrons are multiplied in an avalanche and collected on anodic wires. The signal is read out by capacitive coupling to copper strips, located at both sides of the detector. The strips are arranged in order to be parallel ( $\eta$  coordinate) and orthogonal ( $\phi$  coordinate) to the MDT wires. The time resolution is 1 ns and for this reason the RPC chambers are employed as trigger, while the spatial resolution is 1 cm and for precision measurements MDT are used. TGC are the trigger chambers used in the end-caps. TGC are multi-wire proportional chambers, but unlike the CSC the anode-cathode separation is smaller (1.4 mm) and this results in an improved time resolution. The gas mixture is  $CO_2$  and  $n - C_5H_{12}$  with a percentage of 55% and 45%. The time resolution is 5 ns.

### 2.3.5 The ATLAS trigger system

A great challenge for the LHC experiments is the online event selection, which requires a very efficient trigger system in order to reduce the high rate of events generated during the LHC bunch collisions. The bunch crossing happens each 25 ns, leading to a nominal frequency of 40 MHz. At each bunch crossing around 25 proton-proton interactions happens at the project luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . The amount of data appears to be huge in these conditions and so it's impossible to record all the information: Each event occupies around 1.5 MB, so a recording rate of around 60 TB/s would be needed, but the current technologies allows to record data at about 200 Hz (300 MB/s). A trigger system, able to reduce the event rate by a factor of  $10^6 - 10^7$ , preserving at the same time the interesting events, is of crucial importance. The ATLAS trigger [50] is designed to investigate rapidly the

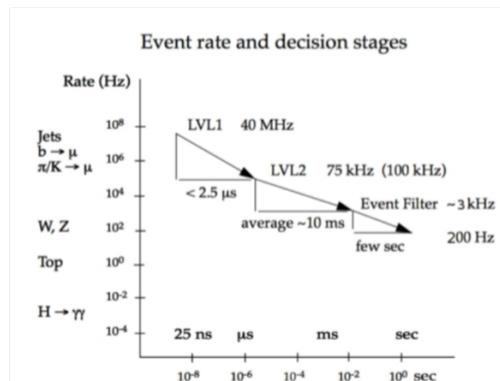


Figure 2.19: Event rate at LHC as function of the processing time. The event rate for rare events of interest is reported on the left.

events detected and choose if record or discard the event, according to the event main features and the compatibility with a set of predefined thresholds contained in the trigger menu. Figure 2.19 reports the event rate at LHC in function of the processing time.

The trigger system was initially composed by a three-level structure: Level-1, Level-2 and Event Filter (EF). Each level refines the measurements of the previous one, applying more stringent selection criteria and combining the information from different detectors.

In Run-2 a two-level trigger system [51] has been implemented in the ATLAS detector, consisting in Level 1 Trigger and High Level Trigger (HLT).

### Level-1 trigger (LVL-1)

The Level-1 trigger uses the information coming from a sub-set of detectors in order to take a decision in less than  $2.5 \mu\text{s}$ , reducing the rate up to 75 kHz. The LVL-1 is completely hardware-based and the detectors used are the calorimeter and the muon spectrometer. The LVL-1 looks for high- $p_T$  muons, electrons, photons, jets and hadronically decay tau-leptons, other than a big amount of missing transverse energy. High- $p_T$  muons are identified with the spectrometer trigger chambers of the barrel and end-cap regions. Calorimetric selections are instead based on information about the granularity. The results coming from the muon trigger and calorimeter are processed by a central processor (CTP).

For each event, the LVL-1 defines one or more *Regions of Interest* (RoI), corresponding to the  $\eta$  and  $\phi$  coordinates of that regions inside the detector where the events of interest have been identified. The data coming from the RoI include informations on the events and the criteria that the event sat-

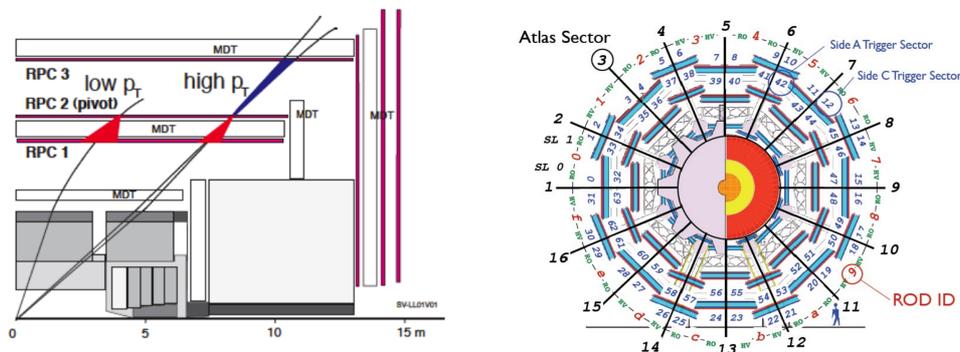


Figure 2.20: Schematic view of the RPC system with low and high- $p_T$  triggers (on the left). x-y plane view of the detector, with 16 sectors of different RPC chambers visible (on the right).

ified (like the threshold passed). These informations are then used by the high-level triggers. A crucial feature of the LVL-1 is the identification of the bunch-crossing of interest without ambiguity.

The Calorimetric Level-1 trigger (L1Calo) is a digital system designed to work with around 7000 trigger towers (TT) ( $0.1 \times 0.1$  in  $\Delta\eta \times \Delta\phi$ ) located in the two calorimeters. It send the results of each bunch crossing to the central processor every  $1.5 \mu\text{s}$ , after the event occurred. Its architecture is relatively compact, with a reduced number of crates and cables, in order to reduce the delay between the event creation and its selection.

The Muon Level-1 trigger works with the RPC and TGC chambers that have a good temporal resolution, so that wrong bunch-crossing identification with muon candidates are avoided. RPC chambers are arranged in 3 concentric double layers: Middle Confirm (RPC1), Middle Pivot (RPC2) and Outer Confirm (RPC3), as can be seen in figure 2.20, and each layer is organized in 16 sectors along the azimuthal coordinate.

The trigger logic is based on the coincidence of hits from different RPC stations, and two different  $p_T$  regimes exist:

- The *low* -  $p_T$  trigger requires a coincidence between the two innermost stations: RPC1 and RPC2. It selects muons with  $p_T$  above the thresholds 4 GeV, 6 GeV and 10 GeV.
- The *high* -  $p_T$  trigger requires the coincidence also of the third external station: RPC3 and selects muons with  $p_T$  above the thresholds 10 GeV, 15 GeV, 20 GeV.

### High-Level Trigger (HLT)

In the Run-2, the Level-2 and the Event Filter have been unified in the High Level trigger (HLT). It's completely software based and takes as inputs the RoIs identified in the LVL-1, refining the measurements using the informations of more detectors. In a first stage, a primary reconstruction/identification of the particles is performed, so that a more precise measure of the parameters is available and more complex selection criteria can be applied respect to the simple LVL-1  $p_T$  thresholds. Then a full reconstruction of the detector, not restricted only to the RoIs, is performed and the algorithms used are mostly the same of the off-line reconstruction algorithms. The final rate is order 1 kHz.

The selected events are stored in CERN computers. These events are common to all the analyses performed in ATLAS; so it's necessary to apply further selections during the off-line analyses, in order to reject the events selected by the trigger but that are not specific for the kind of analysis under study.

### 2.3.6 The Run-3 and the High Luminosity LHC upgrades.

From December 2018 the Run-2 data taking is officially finished. Proton-proton collisions will restart in 2021, with the operations that will be conducted up to the 2023. This period is called Run-3; collisions at center of mass energy of 14 TeV will be performed and the luminosity will reach the value of 2.5 times the nominal one ( $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ ) with expected  $300 \text{ fb}^{-1}$  data collected.

The long shut-down before the start of the Run-3 is called LS2, and in this period new electronics, trigger and detector components are under installation as a first step to prepare the ATLAS detector to the High Luminosity LHC (HL-LHC) that will take place in 2026 (new systems will be also installed during the next long shut-down LS3 in 2024).

The High Luminosity LHC (HL-LHC) is a new phase of the LHC. With upgrades on the accelerator, instantaneous luminosities of a factor 5 larger than the LHC nominal value will be achieved, so that experiments will enlarge their data sample by one order of magnitude compared with the LHC baseline programme. At the HL-LHC will collide beams with about 200 simultaneous collisions per beam crossing; this will increase the ATLAS sensitivity on new or rare physics events, but requires the installation of new performing

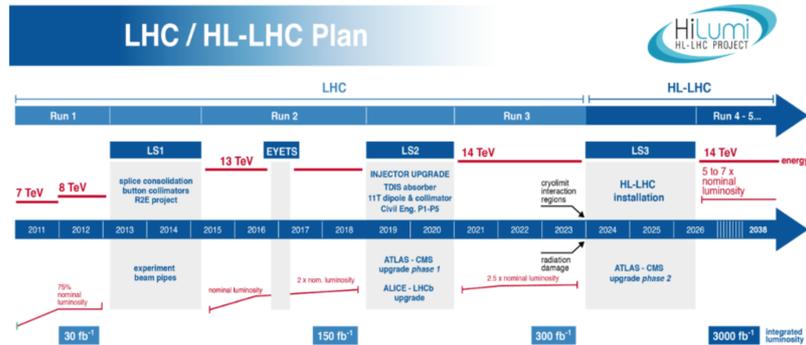


Figure 2.21: LHC timeline.

detectors, able to work in these conditions (hard-radiation, needed of finer granularity and faster readout).

For the Run-3 will be available improvements for the muon spectrometer: *New Small Wheels* that can handle the higher flux of neutrinos and photons expected in the future LHC interactions, 16 new chambers of *Small Monitored Drift Tubes (sMDT)* and *Resistive Plate Chambers (RPCs)* to be installed in the barrel region in order to improve the overall trigger coverage.

There will be improvements also for the Liquid Argon (LAr) calorimeter: new front-end electronics and optical-fibre cabling will be installed, greatly improving the resolution of the detector at trigger level, providing 4 times higher granularity that allow to better differentiate jets of particles from electrons and photons.

# Chapter 3

## Physics objects reconstruction in ATLAS

Physics events from proton-proton collisions selected by the trigger algorithms are recorded for the offline analyses. The full events reconstruction needs sophisticated algorithms developed to provide high efficiency and accuracy in the reconstruction and identification of each physics object.

In this chapter I'll describe the physics objects reconstruction and identification.

### 3.1 Electrons reconstruction and identification

The electrons reconstruction in the detector barrel region ( $|\eta| < 2.47$ ) is performed through both the calorimeter and the Inner Detector informations. The reconstruction algorithm relies on the identification of a cluster in the electromagnetic calorimeter, which is then associated with a track in the Inner Detector.

Seed clusters must have energies greater than 2.5 GeV in a window size of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ , called *tower*. The clusters are seeded from localized energy deposits using a sliding-window algorithm [52]; the center of the seed cluster searches for localized energy deposits, and moves in steps of 0.025 in either the  $\eta$  or  $\phi$  directions, and the process is iterated for every element in the calorimeter. If two seed-cluster candidates are found in proximity, the candidate with the higher transverse energy is retained if its transverse energy is at least 10% higher than the other candidate, otherwise the candidate containing the highest energy central tower is kept.

Charged-particle tracks are reconstructed in the pixel and SCT detectors,

assembling clusters from the hits of the Inner Detector tracking layers. The track reconstruction involves 3 steps: pattern recognition, ambiguity resolution, and TRT extension [53].

Reconstructed tracks are extrapolated to the middle layer of the EM calorimeter and seed clusters are then matched to track candidates; electron-track candidates are refitted using the Gaussian Sum Filter (GSF) [54], to better account for energy loss of charged particles in material, taking into account non-linear bremsstrahlung effects.

In case several GSF-track candidates are matched to the same seed cluster, tracks are ordered according to the minimum distance between the ID and seed cluster coordinates ( $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ ). The electron four-momentum is computed using information from both the primary track and the EM cluster.

Candidates without any track associated are considered as photons.

The forward regions ( $2.5 < |\eta| < 4.9$ ) are not covered by the Inner Detector and the electrons candidates are reconstructed using only information from the electromagnetic calorimeter, clustering nearby cells taking into account the significance of the cell's measured energies with respect to the expected noise. The electron candidates must have a transverse energy greater than 5 GeV and small hadronic energy component. The barycentre of the cluster gives the direction of the electron and the energy is given by the sum of the energies in each cluster cell. It is not possible a distinction between electrons and photons.

In order to avoid objects falsely reconstructed as candidate electrons, it's performed an electron identification [56] in ATLAS by algorithms based on variables with high-discriminating power between the electrons and the other physics objects: longitudinal and lateral shower development, track quality (number of hits), particle identification by the TRT and track-to-cluster matching. Six sets with different quality conditions are defined: *loose*, *loose++*, *medium*, *medium++*, *tight*, *tight++* [57]. They are defined such that tighter categories are a subset of the looser one with higher purity but lower identification efficiency:

- *Loose*: It is based only on calorimeter information. Electron candidates in the  $|\eta| < 2.47$  region, with low hadronic leakage and cuts on shower shape variables (longitudinal and lateral widths) are required, analysing energy deposits in the first and second layer of the ECAL. This category provides high identification efficiency ( $\sim 95\%$ ) but low background rejection (expected rejection of about 500).

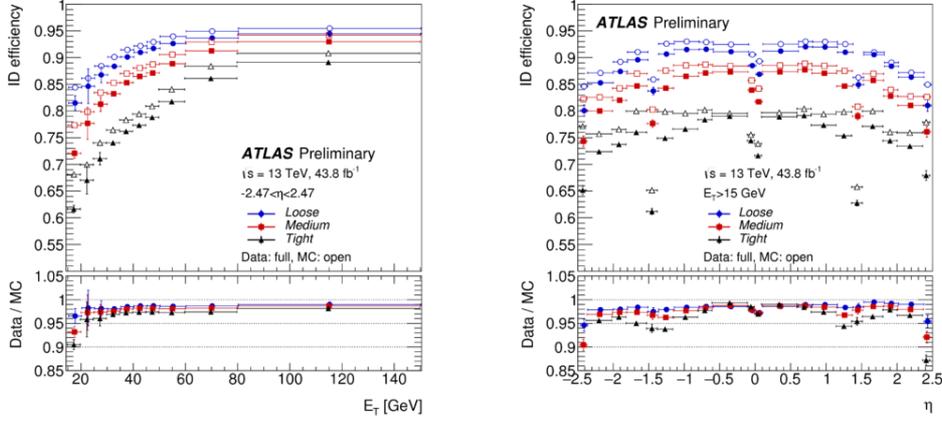


Figure 3.1: Electron identification efficiencies in  $Z \rightarrow ee$  events as a function of transverse energy  $E_T$  (left) and the pseudorapidity  $\eta$  (right), for three operating points based on a likelihood approach: Loose, Medium and Tight. Data efficiencies have been compared with the MC predictions [55].

- *Loose++*: This category adds additional requirements related on the associated track, in order to reject electrons from photons conversion in  $e^+e^-$  couples. It's required at least 1 hit in the pixel detector and at least 7 hits from both the pixel and SCT; moreover the  $\Delta\eta$  between the cluster and the extrapolated track in the first EM layer is required to be smaller than 0.015. The identification efficiency is close to the *Loose* one (93% – 95%) but with an higher expected rejection (of about 5000).
- *Medium*: It adds additional requirements related to the shower shape measured in the first EM layer and requirements about the deviation in the energies of the largest and second largest deposits in this layer, in order to discriminate electrons from  $\pi^0$  and  $\eta$ . Are required additional conditions on the absolute value of the transverse impact parameter of the track and the distance between the cluster and the extrapolated track in the first EM layer:  $|d_0| < 5$  mm,  $\Delta\eta < 0.01$ . The efficiency is of about 88% and the background rejection is higher than the one achieved by the *loose++* selection.
- *Medium++*: It requires at least one hit in the first pixel detector layer to reject electrons from photon conversions. Tracks having a low fraction of high-threshold TRT hits are rejected to decrease the contamination from charged hadrons.  $\Delta\eta$  cluster-track separation in the first

EM layer is required to be less than 0.005. The efficiency is of around 85%, with an expected rejection closer to 50000.

- *Tight*: In this category requirements on separation between the cluster and the matched track, and the absolute value of the transverse impact parameter of the track are tightened:  $\Delta\phi < 0.02$ ,  $\Delta\eta < 0.005$ ,  $|d_0| < 1$  mm; Moreover a requirement on the ratio between the cluster energy and the track momentum ( $E/p$ ) is introduced. The identification efficiency is around 75% with very high background rejection.
- *Tight++*: It adds an asymmetric  $\Delta\phi$  cluster-track separation cut. Efficiency and rejection are slightly better than the *Tight* category.

Also in the forward regions are defined different selections: *loose*, *medium*, *tight*. In order to compensate the absence of informations about tracks, different requirements on the longitudinal and lateral shower shapes are used.

While the identification algorithm used in the Run-I were cut-based, the ones used in the Run-II are likelihood-based (LH) method [58, 59, 60], which relies on the probability density functions of the discriminating variables. A final discriminant based on the likelihoods of the overall signal ( $L_S$ ) and background ( $L_B$ ) is defined:  $\frac{L_S}{L_S+L_B}$ .

Applying the *tag-and-probe* method to  $Z \rightarrow ee$  and  $J/\Psi \rightarrow ee$  datasets [59] it's possible to estimate the electrons reconstruction and identification efficiencies. In figure 3.1 electron identification efficiencies, as a function of transverse energy  $E_T$  and the pseudorapidity  $\eta$ , are shown for 3 operating points based on a likelihood approach: Loose, Medium and Tight. A dataset corresponding to an integrated luminosity of  $43.8 \text{ fb}^{-1}$  recorded by the ATLAS experiment in the 2017, at a centre of mass energy of  $\sqrt{s} = 13 \text{ TeV}$ , was used and the data efficiencies have been compared with the MC predictions. The lower data efficiencies than the MC ones arises from mismodelling of calorimeter shower shapes in the GEANT4 detector simulation.

To suppress objects that are falsely reconstructed as prompt leptons, e.g. from semileptonic heavy-flavour hadron decays or misidentified jets, the electrons in the event are usually required to be isolated. Calorimeter isolation, track isolation or both criteria can be applied. They exploit the total deposited transverse energy of topological clusters and the sum of the transverse momenta of all tracks in a region around the identified electron, for the calorimeter and track isolation respectively. The selection of

the tracks relies on some basic quality cuts, like number of hits and requirements on the longitudinal impact parameter  $z_0$ . The region investigated is defined by a cone of a given radius  $R$ , and the expected contribution of the electron itself is excluded.

All the electron isolation working points used in ATLAS are summarized in table 3.1. The *LooseTrackOnly* working point is optimised in order to provide high signal efficiencies (around 99% on average), almost constant in  $\eta$  and  $p_T$ , but with a related low background rejection (order of 15).

The *LooseTrackOnly*, and all the working points relying on track isolation, are based on a variable-cone denoted as  $ptvarcone(R_{max})$ , where the cone size gets smaller for larger transverse momentum of the electron:  $R = \min(R_{max}, 10\text{GeV}/p_T)$ . In very busy environments or for boosted particles, other objects can end up very close to the electron candidates, so  $ptvarcone$  has higher signal efficiency respect to the fixed-cone variable ( $ptcone$ ). The *LooseTrackOnly* working point is defined by setting  $R = 0.2$  and varying the requirements imposed to  $ptvarcone(R)/p_T$  in order to have a constant efficiency in  $\eta$  and  $p_T$ .

Working Point	Calo Isolation	Track Isolation
LooseTrackOnly	-	99%
FixedCutTight	Cut: $topoetcone20/p_T < 0.06$	Cut: $ptvarcone20/p_T < 0.06$
FixedCutLoose	Cut: $topoetcone20/p_T < 0.2$	Cut: $ptvarcone20/p_T < 0.15$
FixedCutTightTrackOnly	-	Cut: $ptvarcone20/p_T < 0.06$
FixedCutHighPtCaloOnly	Cut: $topoetcone20 < 3.5 \text{ GeV}$	-

Table 3.1: Electron isolation working points. The  $topoetcone20$  variable exploits the total deposited transverse energy of topological clusters around the identified electron in a cone of radius  $R = 0.2$ . The  $ptvarcone20$  variable exploits the sum of the transverse momenta of all tracks around the identified electron in a variable-size cone of radius  $R = \min(0.2, 10\text{GeV}/p_T)$ .

## 3.2 Muons reconstruction and identification

Muons are reconstructed through the informations of the Muon Spectrometer and of the Inner Detector. Indeed, muons originating from hard-scatter events produce hits in the inner tracking detectors, pass through the calorimeters depositing only a minimal fraction of the initial energy, and then reach the Muon Spectrometer.

The track reconstruction in the Muon Spectrometer proceeds in four steps [61]: search for hit patterns inside each muon chamber, in order to make segments; in each MDT and nearby trigger chambers, track segments are built by straight-line fit to the hits found in each layer, in the bending plane of the detector; muon track candidates are built by fitting hits combining segments of different layers. At least two matching segments are required to build a track, except in the barrel–endcap transition region, where a single high-quality segment is sufficient to build a track. The same segment can be used to build several track candidates in a first moment, and then an overlap removal algorithm selects the best assignment. As last step, the hits associated with each track candidate are fitted using a global  $\chi^2$  fit. Hits providing large contributions to the  $\chi^2$  are removed and the track fit is repeated. Also an hit recovery procedure is performed, looking for additional hits consistent with the candidate track trajectory, refitting if additional hits are found.

Different techniques have been implemented for the muon reconstruction and identification in the ATLAS detector. These strategies exploit the technology of the different sub-detectors, that provides complementary methods for the muon track reconstruction [61]:

- *Combined (CB) muons*: Tracks are separately reconstructed in the Muon Spectrometer and in the Inner Detector in a first stage, and then are combined with a global refit that uses the hits from both the two sub-detectors. Combined reconstruction improves the momentum resolution and the rejection of muons coming from pions and kaons decays. Due to the ID coverage, combined muons are defined in the region  $|\eta| < 2.5$ .
- *Segment tagged (ST) muons*: Low  $p_T$  muons could not cross so many stations for reconstructing a track in the Muon Spectrometer and so to be reconstructed as *combined* muon. *ST* are low  $p_T$  muons with the ID track associated with at least one local track segment in the first layer of the spectrometer. *ST* are used also for muons that enter in regions with reduced MS acceptance (like support structures and service passages).
- *Calorimeter Tagged (CaloTag) muons*: They are muons with an ID track extrapolated to the calorimeters and matched to an energy deposit consistent with a minimum-ionizing particle. *CaloTag* muons have the lowest purity respect to all the other muon types, but they are used to recover efficiency in the region where the Muon Spectrometer is only partially instrumented ( $|\eta| \sim 0$ ).

- *Extrapolated (ME) muons*: If no track in the ID is found, muons are reconstructed using only the Muon Spectrometer with loose requirement on the compatibility with originating from the Interaction Point. *ME* muons are mainly used to extend the muon reconstruction acceptance in the detector region not covered by the ID ( $2.5 < |\eta| < 2.7$ ).

As for the electron candidates, it's performed a muons identification by the application of quality criteria to suppress background (mainly from the light hadron decays) and select prompt muons. Different discriminating variables between prompt muons and background muons are used to perform the identification. For *combined* muons, in addition to requirements on track quality, like the number of hits in the Inner Detector and the Muon Spectrometer, criteria related to the compatibility of the ID and MS measurements are investigated:

- $\Delta(q/p)$  *significance*: Absolute value of the difference between the ratio of the charge and momentum of the muons measured in the ID and MS. This quantity is divided by the corresponding uncertainty on the difference;
- $\rho' = \Delta p_T/p_T$ : Absolute value of the difference between the transverse momentum measured in the ID and in the MS, divided by the combined track transverse momentum;
- Normalized  $\chi^2$  of the combined track fit.

4 muon identification selections, that provide different background rejections and signal efficiencies, are defined:

- *Medium Muons*: This is the standard identification criterion in ATLAS, optimised to minimise the systematic uncertainties associated with the reconstruction and calibration of the muons candidates. In this category only the CB and ME tracks are used, with hits requirements in the MDT layers for CB muons and in the MDT/CSC layers for ME. The matching between the Inner Detector and the Muon Spectrometer must satisfy a loose criteria ( $\Delta(q/p)$  significance  $< 7$ ) to suppress the contamination of hadrons misidentified as muons.
- *Loose Muons*: These muons are mainly used to reconstruct Higgs boson candidates in the 4-lepton final state, selecting all muon types: in the region  $|\eta| < 2.5$  about 97.5% of *loose* muons are CB muons,  $\sim 1.5\%$  are CT and  $\sim 1\%$  are ST muons.

- *Tight Muons*: This is the category with the highest purity. Only CB muons are used and hits in at least two stations of the MS, that satisfy the medium muons selection, are required. The criteria related to the compatibility of the ID and MS measurements are: normalized  $\chi^2$  of the combined track fit less than 8 to remove pathological tracks, and to maximize the background rejection for momenta below 20 GeV, where the misidentification probability is higher, a two dimensional cut on  $\Delta(q/p)$  significance and  $\rho'$  as a function of the muon  $p_T$  is performed.
- *High  $p_T$  Muons*: This selection is defined in order to maximise the momentum resolution for muons with transverse momentum above 100 GeV. The selection is optimised for searches for high-mass  $Z'$  and  $W'$  resonances. Are used CB muons that satisfy the medium selection and having at least three hits in three MS stations. Requiring three MS stations reduces the reconstruction efficiency by about 20%, but improves the  $p_T$  resolution of above 1.5 TeV by approximately 30%.

The reconstruction efficiency for the muons can be estimated using the *tag-and-probe* method on  $Z \rightarrow \mu\mu$  data samples, and it is divided in two steps: first the identification efficiency, assuming a reconstructed track in the Inner Detector, is measured using calorimeter-tagged muons as probes; then the efficiency is corrected taking into account the ID-track reconstruction efficiency, using standalone muons as probes:  $\epsilon(X) = \epsilon(X|CT) \cdot \epsilon(ID|MS)$ , where  $X = \textit{loose, medium, tight or High-}p_T$ . This is based on the assumption that the ID and MS track reconstructions are independent from each other. Figure 3.2 shows the muon reconstruction efficiencies, as a function of  $\eta$ , for the different muon selections. Loose and Medium selections have very similar efficiencies in the  $|\eta| < 2.5$  range, with the exception of the region  $|\eta| < 0.1$ , where the Loose selection fills the MS acceptance gap using the CT and ST muons contributions. The efficiencies of Loose and Medium selections are observed to be above of 98%, and between 90 and 98% for the Tight selection, with very good agreement with the efficiencies predicted by the simulation. The lower values of the *High- $p_T$*  selection are mainly due to the tight requirement on momentum resolution.

Isolation criteria similar to the ones applied for the electrons are required, in order to select muons originating from  $W$  and  $Z$  decays, and reject muons from hadronic decays. All the muon isolation working points used in ATLAS are summarized in table 3.2.

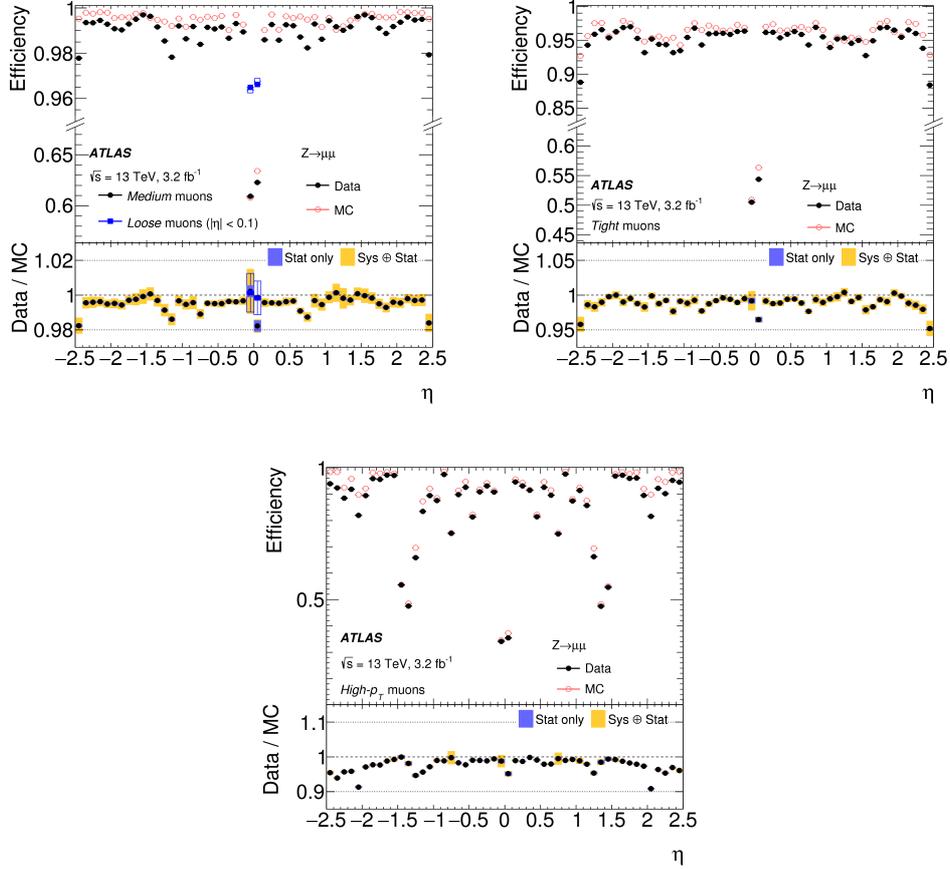


Figure 3.2: Muon reconstruction efficiency, as a function of  $\eta$ , measured on  $Z \rightarrow \mu\mu$  events for Medium (upper left), Tight (upper right) and High- $p_T$  (lower) selections [61].

### 3.3 Jets reconstruction

Due to the color confinement of the QCD, quarks and gluons cannot be observed directly. They can't exist in free form but hadronise almost immediately after they are produced: quarks create other couples of quarks, that strongly interact with each other and create (colour-neutral) hadrons. The hadrons produced in this process can be narrowed to a cone, referred to as jet. For this reason, when quarks are produced in any interaction, it is not possible to see the tracks left by them, but the signature is a big amount of tracks due to the hadrons coming from initial quarks.

Working Point	Calo Isolation	Track Isolation
LooseTrackOnly	-	99%
FixedCutTight	Cut: $topoetcone20/p_T < 0.06$	Cut: $ptvarcone30/p_T < 0.06$
FixedCutLoose	Cut: $topoetcone20/p_T < 0.2$	Cut: $ptvarcone30/p_T < 0.15$
FixedCutTightTrackOnly	-	Cut: $ptvarcone30/p_T < 0.06$
FixedCutHighPtTrackOnly	-	Cut: $ptcone20 < 1.25$ GeV

Table 3.2: Muon isolation working points. The  $topoetcone20$  variable exploits the total deposited transverse energy of topological clusters around the identified muon in a cone of radius  $R = 0.2$ . The  $ptcone20$  and  $ptvarcone30$  variables exploit the sum of the transverse momenta of all tracks around the identified muon in a fixed cone of radius  $R = 0.2$  and in a variable-size cone of radius  $R = \min(0.3, 10\text{GeV}/p_T)$ , respectively.

Jet algorithms used in the reconstruction provide rules for grouping particles into jets, involving distance measures and recombination schemes, in order to assign a momentum to the resulting jet that reflects the properties of the original parton.

Hadronic jets are reconstructed starting from calorimeters clusters. Particles usually deposit their energy in many calorimetric cells, both in lateral and longitudinal directions. Clustering algorithms [62] are designed with the aim of group these cells and sum the total energy deposit inside each cluster, according to certain particular criteria. These clusters are called topological cluster or *topocluster*. The algorithm identifies a seed cell with a signal-to-noise ratio ( $|E_{cell}/\sigma_{noise}|$ ) greater than 4.  $E_{cell}$  is the cell energy and the noise is the RMS of the energy distribution in events triggered at random bunch crossing. Neighbouring cells around the seed are included if  $|E_{cell}/\sigma_{noise}| > 2$ , and finally added to form the topological cluster. Topoclusters are defined as massless and their final energy correspond to the sum of the energies of all cells included. The direction is calculated from weighted averages of  $\eta$  and  $\phi$  of the constituent cells, where the weight is the absolute cell energy. Because of calorimeter noise fluctuations, clusters can have a negative energy. When this happens, energy clusters are rejected entirely from the jet reconstruction since they do not have physical meaning.

In order to provide a good description of the QCD processes, that take place in the collisions, a good jet reconstruction algorithm has to guarantee the following properties:

- Collinear safety: the split of one particle into two collinear particles

has no effect on the jet reconstruction;

- Infra-red safety: the presence of additional soft particles between jet components does not affect the jet reconstruction;
- Resolution and other detector effects should have a minimal impact on the jet reconstruction;
- Invariance under Lorentz boosts along beam axis is required;
- Minimum computer resources used.

Among the different jet reconstruction algorithms, I'll describe those used in this thesis work: *anti* –  $k_T$  and Track Calo Cluster jets.

### The anti – $k_T$ algorithm

ATLAS uses the *anti* –  $k_T$  algorithm [63] for the jet reconstruction. The algorithm is based on the definition of the distance  $d_{ij}$  between two objects  $i$  and  $j$ , and the distance  $d_{iB}$  between the  $i_{th}$  object and the beam:

$$d_{ij} = \min(k_{T_i}^{2p}, k_{T_j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.1)$$

$$d_{iB} = k_{T_i}^{2p} \quad (3.2)$$

where  $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$  is the angular distance between the two objects and  $k_{T_i}$ ,  $y_i$  and  $\phi_i$  are the transverse momentum, the rapidity and the azimuthal angle of the  $i_{th}$  object.  $R$  and  $p$  are parameters of the algorithm. The clustering proceeds in steps:

- For each particle  $i$ , the two distances  $d_{ij}$  and  $d_{iB}$  are computed;
- the algorithm calculates  $\min(d_{ij}, d_{iB})$ ;
- if  $\min(d_{ij}, d_{iB}) = d_{iB}$ , the input  $i$  is considered to form a jet and it is removed from the list of inputs;
- else if  $\min(d_{ij}, d_i) = d_{ij}$ , the inputs  $i$  and  $j$  are combined into one single input, using the sum of four-momentum of each input, and the resulting object is given into the list of possible inputs, while  $i$  and  $j$  are removed.

The procedure is repeated until no inputs are left, so that all inputs in the event will end in a jet. Different values of the parameter  $p$  leads to different algorithms:

- $p = -1$ : *anti* –  $k_T$  algorithm [63];
- $p = 0$ : Cambridge/Aachen algorithm [64];
- $p = 1$ :  $k_T$  algorithm [65].

For the *anti* –  $k_T$  algorithm the distances are given by:

$$d_{ij} = \min(1/k_{T_i}^2, 1/k_{T_j}^2) \frac{\Delta R_{ij}^2}{R^2} \quad (3.3)$$

$$d_{iB} = 1/k_{T_i}^2 \quad (3.4)$$

The distance is therefore dominated by the particle with the greatest transverse momentum, hence soft inputs will cluster preferentially with hard inputs, instead of clustering with other soft particles. So a jet is defined around the particle with the highest  $k_T$ , and if the hard particle has no hard neighbours within a distance  $2R$ , then it will simply accumulate all the soft particles within a circle of radius  $R$ , resulting in a perfectly conical jet.

Instead all hard inputs with  $\Delta R_{ij} < R$  will be combined into one jet, ensuring the collinear safety. If  $k_{T_i} \gg k_{T_j}$ , the jet will be conical and centred on  $k_{T_i}$ . For  $k_{T_i} \sim k_{T_j}$ , the shape will instead be the union of cones (radius  $< R$ ) around each hard particle plus a cone (of radius  $R$ ) centred on the final jet. If two hard particles are present with a distance  $R < \Delta R_{ij} < 2R$ , there will be two hard jets. If  $k_{T_i} \gg k_{T_j}$ , the jet around the hard  $i$ -particle will be conical, and the other jet will be partially conical due to the missing part overlapped with the most energetic jet. Instead if  $k_{T_i} \sim k_{T_j}$  neither jets will be conical and the overlapping part is given by  $\Delta R_{ib}/k_{T_i} = \Delta R_{jb}/k_{T_j}$ . In figure 3.3 clustering with the *anti* –  $k_T$  algorithm are showed.

Different kind of jets are defined depending on the radius parameter  $R$ : *Small- $R$  jets* are reconstructed from topo-clusters at the electromagnetic scale (EM scale) with  $R = 0.4$ , while *Large- $R$  jets* are reconstructed from topo-clusters at the Local Cluster Weighting Calibration (LCW) scale with  $R = 1.0$ .

The impact of soft radiation, coming from underlying events or pile-up interactions, on the measurement of the jet properties is mitigated using two different approaches for small- $R$  and large- $R$  jets.

In small- $R$  jets the *jet-vertex tagging* (JVT) technique [66] is adopted for the pile-up suppression. It employs a multivariate combination of two variables:

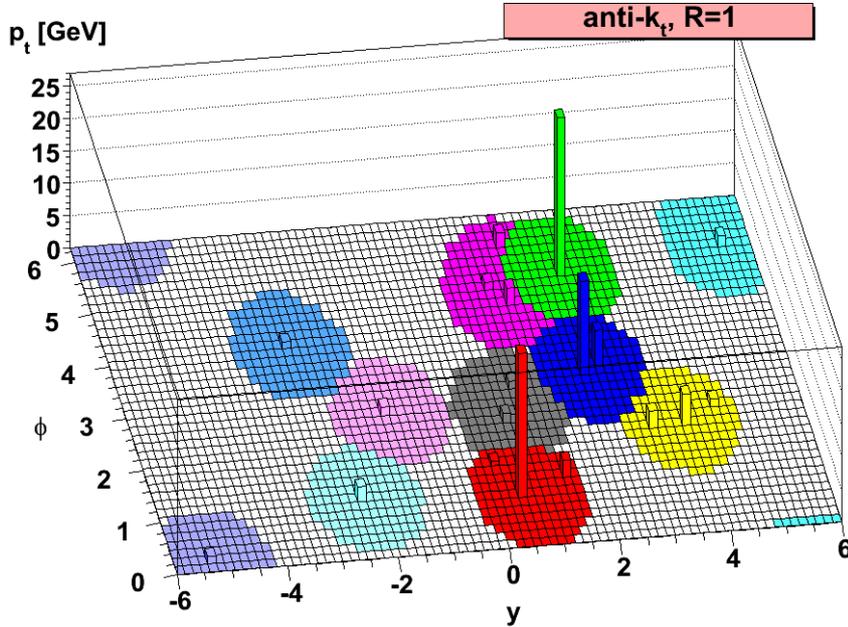


Figure 3.3: Shapes of jets in the calorimeter, reconstructed with the  $anti-k_T$  algorithm [63].

- *Jet Vertex Fraction* (JVF), that is the ratio of the scalar sum of the tracks momentum inside the jet originate from the hard-scatter vertex, and the sum of the momentum of the tracks inside the jets from any primary vertex.
- $R_{p_T}$ , defined as the ratio of the scalar sum of the tracks momentum inside the jet originate from the hard-scatter vertex, and the  $p_T$  of the fully calibrated jet that includes the pileup subtraction.

These variables have a range from 0 to 1 and pile-up jets tend to have smaller values and peak at 0, while hard-scatter jets larger values and peak at unity, as it is possible to see in figure 3.4, where the correlation plot for these two variable is showed for pile-up jets and hard-scatter jets.

In Large-R jets, pileup and soft radiations are individually identified and removed, by resolving the substructure of the hard large-R jet. This approach is called *grooming* [67] and the algorithm used in this thesis is the *trimming* [68] one: the constituents of a large-R jet are re-clustered using the  $k_T$  algorithm with radius parameter  $R = 0.2$ ; the transverse momentum of each subset  $i$  is evaluated and if the fraction of transverse momentum is

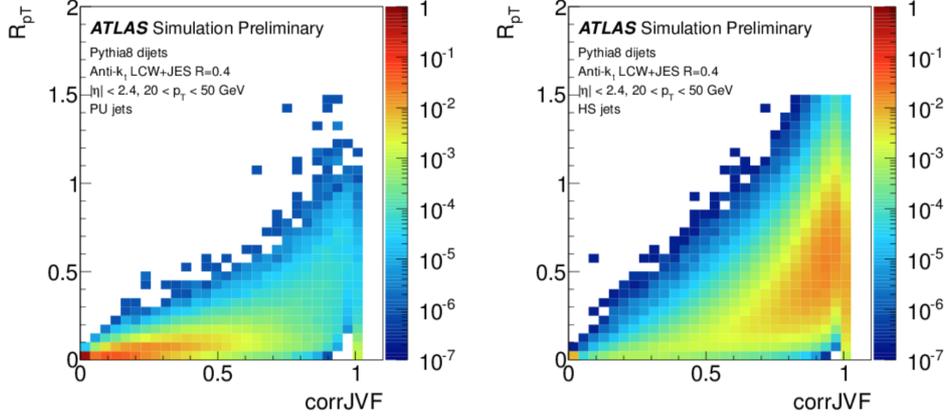


Figure 3.4: 2-dimensional correlation of JVT and  $R_{pT}$  for pile-up (left) and hard scatter (right) jets [66].

lower than 5% (i.e.  $p_T^i/p_T^{jet} < 0.05$ ), the subset is discarded and the remaining subsets are assembled to form the trimmed large-R jet. Figure 3.5 shows an illustration of the trimming procedure.

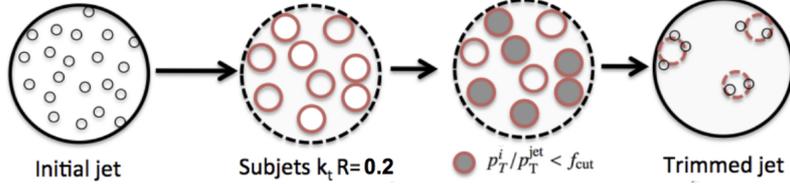


Figure 3.5: Illustration of the trimming procedure for pileup and soft radiations in large-R jets.

### Track Calo Cluster Jets (TCC Jets)

Jet substructure techniques are very important in ATLAS for searches for new physics. The study of resolvable energy structures within a jet is a powerful way to discriminate between hadronic decays of massive particles and QCD multi-jet events. As an example, the decay of a vectorial boson, like  $W$  or  $Z$ , into a pair of quarks generally produces two significant energy deposit regions in the detector, while a single quark or gluon from a QCD multi-jet event produces one significant energy deposit region. As the momentum of the vector boson increased, the quarks pair is increasingly boosted, and so they are increasingly collimated until they are reconstructed as a single

large-R jet. So the jet substructure it's important in order to differentiate the quarks pair from the single quark or gluon.

This makes use of the energy and angular resolution of the detector employed to reconstruct the jet. ATLAS has mainly focused on the use of calorimeter-based jet substructure, in order to takes advantage of the exceptional energy resolution of the ATLAS calorimetry. But for very energetic events, jets could be so collimated that the calorimeter lacks the angular resolution to resolve the desired structure within the jet. The tracking detectors instead have excellent angular resolution and good reconstruction efficiency at very high energy, while their energy resolution deteriorates.

In this thesis new unified objects, built from both tracking and calorimeter informations, the Track-CaloClusters (TCCs) [69], are used for large-R jet reconstruction. TCC algorithm employs the spatial coordinates of the tracker and the energy scale of the calorimeter, correlating low-level objects (tracks and individual energy deposits in the calorimeter) before running any jet finding algorithms. The resulting objects are used as inputs to jet reconstruction, and this results in improved resolution for both jet mass and substructure variables.

The TCC matching procedure try to match every good quality track to every topo-cluster. The uncertainty on the track extrapolation to the calorimeter is compared to the width of the topo-cluster and, if this uncertainty is larger than the topo-cluster width, the track is discarded from the matching procedure, otherwise the track-cluster pair is defined as matched if their angular separation satisfy the condition  $\Delta R < \sqrt{\sigma_{cluster}^2 + \sigma_{track}^2}$ , where  $\sigma_{cluster}$  is the topo-cluster width and  $\sigma_{track}$  the uncertainty on the track extrapolation.

The matching procedure makes use of tracks originating from any primary vertex, not only of tracks from the selected hard scatter primary vertex. The tracking detector is able to differentiate between multiple in-time collisions, while the calorimeter cannot and thus tracks from pile-up vertices matched to topo-clusters are not used to build TCC objects. Three categories of TCC objects are defined:

- *Combined TCCs*: Tracks compatible with the primary vertex and matched to topo-clusters.
- *Neutral TCCs*: Topo-clusters not matched to tracks from any vertex.
- *Charged TCCs*: Tracks from the primary vertex not matched to clusters.

In figure 3.6, the resulting fractions of neutral, charged, and combined TCCs are showed for simulated  $W' \rightarrow WZ \rightarrow qq\bar{q}\bar{q}$  events, as a function of the TCC  $p_T$ . There is a large number of neutral TCCs in the *low*  $p_T$  range, while the number of charged TCCs is small on the whole  $p_T$  range. The combined TCC becomes the most common scenario for  $p_T \geq 5$  GeV, as expected since highly energetic topo-clusters are likely from the hadronically decaying  $W$  or  $Z$  bosons, and thus close to tracks.

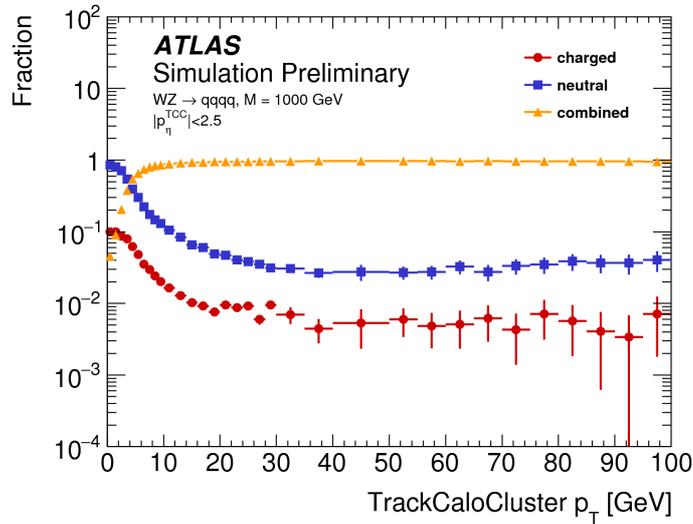


Figure 3.6: Fractions of different Track-CaloCluster classes as a function of the TCC  $p_T$  [69].

If there is a match between multiple tracks and a single cluster, multiple topo-clusters and a single track, or multiple topo-clusters with multiple tracks, the TCC reconstruction procedure still creates exactly one TCC object per track originating from the primary vertex, where the track angular coordinates are used, but the scale coordinates must be adapted to account for energy sharing between the different matches:

- Multiple clusters can be matched to the same track. Each cluster matched to the seed track should contribute to the resulting TCC object proportionally to its  $p_T$  fraction, respect to all the matched clusters.
- Each cluster may contribute to multiple TCC objects, so multiple tracks are matched to the same topo-cluster. The cluster contribution has to be weighted by the fraction of track  $p_T$ , respect to all the matched tracks.

With this method a splitting of the energies is performed without requiring a precise measurement of the track  $p_T$ , since only ratios with respect to the  $p_T$  of the sum of all matching tracks is required, and so the algorithm is less sensitive to the scale of any particular track. The main thing is whether or not any individual track represents a significant fraction of the total  $p_T$ . Moreover the tracker and calorimeter energy measurements are never directly compared.

### Jet Energy Scale (JES) e Jet Energy Resolution (JER)

Reconstructed jets do not have energies corresponding to the effective “real” ones since different effects need to be taken into account:

- *Calorimeter non-compensation*: Partial energy measurement deposited by hadrons;
- *Passive material*: Loss of energy in the detector passive material;
- *Leakage*: Loss of energy due to particles of the hadronic shower that end up over the acceptance of the calorimeter;
- *Out-of-cone leakage*: Energy deposits of particles belonging to the jet but not included in the reconstructed jet;
- *Noise thresholds and jet reconstruction inefficiencies*: Signal losses in the clustering process due to threshold effects and jet reconstruction inefficiencies.

The energy measured in the calorimeter clusters so needs to be calibrated in order to assign the correct value. This procedure is known as Jet Energy Scale (JES) [70], and consists in different steps.

The first step is the reconstruction of the jet energy at the electromagnetic scale, that is the basic calorimeter signal scale for the ATLAS calorimeters. A signal from the calorimeter is first calibrated as the signal coming from an electron. The EM scale is obtained using measurements of electrons taken during the test-beams both in the barrel and in the endcap calorimeters [71, 72, 73], and has been validated using muons signals coming from the test-beams and cosmic-rays.

This EM scale calibration provides a very good description for energy deposits produced by electrons and photons, but not for deposits due to hadronic particles, and for this reason jets need to be converted from the EM scale to the hadronic scale. In ATLAS several calibration schemes, with different

levels of complexity and different sensitivity to systematic effects, have been implemented [74]:

- *EM+JES calibration*: This is a Monte Carlo-derived jet calibration, which applies a simple jet-by-jet correction. Each jet at the EM scale is scaled by a correction factor, which is a function of the reconstructed jet energy and  $\eta$ . Other than the energy correction, also pile-up and a jet origin correction are applied;
- *Global Sequential (GS) Calibration* [74]: This is a Monte Carlo-derived jet calibration, which applies EM+JES calibration to the jets and then they are scaled by a jet-by-jet correction factor, which depends on the jet  $p_T$ ,  $\eta$  and several longitudinal and transverse jet properties. The main effect is to improve the jet energy resolution (JER);
- *Global Cell Energy-Density Weighting Calibration (GCW)* [75]: The goal of this calibration is to compensate for different calorimeters' responses to hadrons and electromagnetic particles, by weighting each cell of the jet cluster. The weights are determined by minimizing the energy fluctuations between the reconstructed jets and particle jets in Monte Carlo simulation. Jets at the EM scale are weighted and a final jet energy scale correction is applied;
- *Local Cluster Weighting Calibration (LCW)*: This calibration uses properties of topoclusters, like their energy, depth in the calorimeter, cell energy density, fractional energy deposited in the calorimeter layer and energy measured around it, to calibrate them individually before applying jet reconstruction. These weights are determined from Monte Carlo simulations of charged and neutral pions. In the same way of the GCW scheme, a final correction of the jet energy is applied.

The width of the reconstructed energy distribution provides the *Jet Energy Resolution* (JER). Deviation in the jets response could arise from the stochastic nature of the hadronic shower, calorimetric electronic noise and pile-up effects.

In events containing only two jets, the  $p_T$  of the two jets shall be balanced because of the momentum conservation in the transverse plane. JER can be measured studying the asymmetry observed between the jet  $p_T$  in the same pseudorapidity region (in order to minimize detector effects). The energy resolution is obtained in  $p_T \times \eta$  bins and can be parametrised as follows [76]:

$$\frac{\sigma_{p_T}}{p_T} = \frac{N}{p_T} \oplus \frac{S}{\sqrt{p_T}} \oplus c \quad (3.5)$$

Where  $N$ ,  $S$ , and  $c$  are the noise, stochastic and constant terms respectively. The measurement is performed with the method described above and the distribution of the results can be build for each  $\eta$  bin. Then a fit using the functional form of the equation 3.5 can be done. From equation 3.5 it is possible to see that JER becomes less significant for high jet energies.

## Jet quality

In order to have an handle to suppress fake-jets that could be originated in non-collision background processes (beam gas, halo events, cosmic-ray muons) or due to noisy channels of the calorimeter and its electronics, in ATLAS there are different quality criteria for the reconstruction of jets: *Looser*, *Loose*, *Medium* and *Tight* [77].

Several variables are exploited to distinguish between ‘good’ and ‘bad’ jets:

- Signal shape in the LAr calorimeters;
- Shower-development via energy ratios;
- Charged-particle contributions from tracking information.

The 4 categories have different ‘good’ jet selection efficiencies and ‘bad’ jet rejections. The *Looser* selection is designed to provide an efficiency above 99.8%, with a rejection factor of about 50%. It rejects most of the fake jets due to calorimeter noise; the *Tight* selection has lower efficiencies but the highest fake rate rejection.

## 3.4 b-tagging

The b-tagging identification is a crucial feature in top quark reconstruction and in b-tagged analyses. The b-hadrons are particles with a lifetime relatively long (around  $1 \times 10^{-12}$  s) and they can travel up to 3 mm from the primary vertex before decaying. Tracks from b-hadrons will point to a secondary vertex and will be characterized by big impact parameters respect to the primary vertex. In figure 3.7 is illustrated a jet with a secondary vertex reconstructed from displaced tracks, with large impact parameter and with a significant decay length, indicating the decay of a heavy, long-lived particle (b- or c- hadron).

The capacity of ATLAS to identify and tag b-jets has been enhanced in the Run-2 with the addition of the Insertable B Layer (IBL) [78], and improvements in the tracking and b-tagging algorithms led to an improvement in the

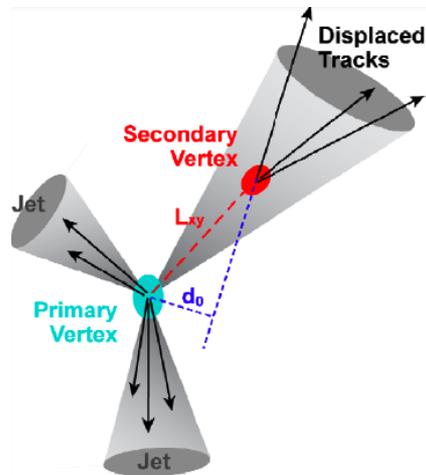


Figure 3.7: Illustration of a jet with a secondary vertex reconstructed from displaced tracks, with large impact parameter and with a significant decay length.

low and medium jet  $p_T$  region.

Different algorithms have been developed for the jet b-tag identification. The tagging is applied to jet reconstructed with the *anti* -  $k_T$  algorithm, with  $p_T > 20$  GeV and  $|\eta| < 2.5$ . The JVT algorithm for the pile-up suppression is also applied.

The inputs to the b-tagging algorithms are the tracks of the charged particles reconstructed in the Inner Detector ( $|\eta| < 2.5$  region). The tracks are matched to the jets and then are required to pass some quality requirements. Three different algorithms, that provides complementary informations, are used in ATLAS [79]:

- **IP2D, IP3D** are algorithms based on the longitudinal and transverse impact parameters;
- **SV**, a vertex-based algorithm that attempts to reconstruct the secondary vertex;
- **JetFitter**, that is based on the reconstruction of the expected topological structure of the b- and c-hadron decay chain.

The outputs of these algorithms are combined in a multivariate discriminant, **MV2 (MultiVariate Tagger)**, that provides the best separation between b-jets from light (u, d, s-jets or gluon jets) and c-jets.

The tracks are matched to the calorimetric jets depending on their angular

separation  $\Delta R(track, jet)$  with a variable matching value as a function of the jet  $p_T$ , resulting in a narrower cone for jets with high  $p_T$ , that are more collimated. A given track is matched with solely one jet and, if the track can be associated with more jets, the jet that gives the minimum  $\Delta R$  value is chosen.

The selection criteria instead are variable and depends on the b-tagging algorithm. IP2D and IP3D have a more stringent criteria. The relevant requirements are: tracks  $p_T > 1$  GeV, transverse and longitudinal impact parameters  $|d_0| < 1$  mm,  $|z_0 \times \sin\theta| < 1.5$  mm and at least two hits in the Pixel Detector. SV algorithm applies a less stringent selection: tracks  $p_T$  are required to be greater than 700-800 MeV and less significant requirement are imposed on the impact parameters and the quality of the tracks.

### IP2D, IP3D algorithms

Due to their long lifetimes, b-hadron decay vertexes are displaced from primary vertex. The transverse impact parameter  $d_0$  and the longitudinal impact parameter  $|z_0 \times \sin\theta|$  measure this distance. The IP2D and IP3D algorithms make use of the significance parameters,  $d_0/\sigma_{d_0}$  and  $z_0\sin\theta/\sigma_{z_0\sin\theta}$ , to discriminate tracks matched to jets.

A sign is assigned to the impact parameter, as function of primary vertex and the jet direction: positive if primary vertex is in front with respect to the jet direction, negative in case the vertex is behind the jet direction. Both transverse and longitudinal impact parameters informations are used in IP3D, while IP2D uses only transverse impact parameters (and it results in a discriminant not depending on pile-up effects, which are related to the longitudinal impact parameter). A log-likelihood-ratio (LLR) discriminant is built for b-jets, c-jets and light jets separations. Figure 3.8 shows the distributions for  $d_0/\sigma_{d_0}$  and  $z_0\sin\theta/\sigma_{z_0\sin\theta}$  for b-, c- and light jets.

### SV algorithm

The secondary vertex algorithm has the goal of the reconstruction of the displaced secondary vertex within the jet. Candidate tracks are considered in pairs and used to reconstruct two-track vertexes. Candidate pairs tracks are rejected if they form a secondary vertex that can be associated to a long-lived particle ( $K_S$  or  $\Lambda$ ), photon conversion or hadronic interaction with the detector material. In figure 3.9 is showed the final reconstruction efficiency, as function of jet  $p_T$  and  $\eta$  for b-, c- and light jets.

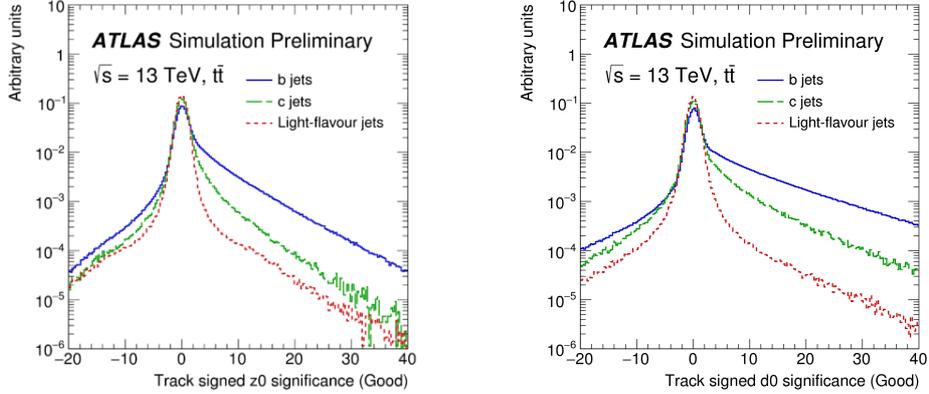


Figure 3.8: The transverse (left) and longitudinal (right) signed impact parameter significance of the tracks for b (blue line), c (green line) and light-flavour (red line) jets in  $t\bar{t}$  events [79].

### JetFitter algorithm

This algorithm reconstruct the whole decay chain  $PV \rightarrow b \rightarrow c$ , exploiting the expected topological structure of the weak b-hadron decays and c-hadron decays. A Kalman filter is used in order to find a common line on which the primary, bottom and charm vertices lie. In figure 3.10 some output variables distributions are shown.

### MV2 multivariate algorithm

The outputs of the three algorithms described above are combined in a *boosted decision tree* (BDT) algorithm, in order to discriminate the b-jets from the light and c-jets.

The training of the multivariate algorithm is performed on a dataset of 5 millions of  $t\bar{t}$  events. In the training also the kinematic of the jets ( $p_T$  and  $\eta$ ) is taken into account, in order to exploit correlations with the other input variables. In order to avoid any difference of the kinematic spectra between signal and background jets, that could be interpreted as discriminant by the training, the signal jets are reweighed to match the spectrum of the background jets. The MV2cXX algorithm is defined as the output of the BDT, with the training performed considering the b-jets as signal and the background composed by light and c-jets.

Three different versions of the algorithm are implemented: MV2c00, MV2c10 and MV2c20, where cXX represents the c-jet fraction in the training: in

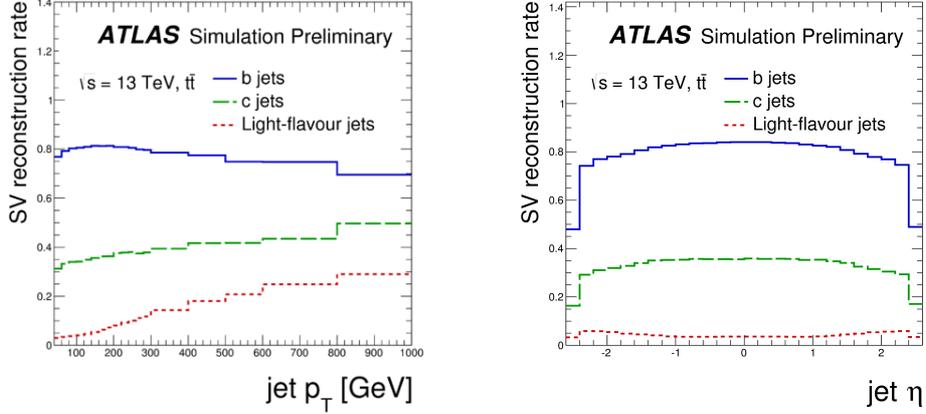


Figure 3.9: Secondary vertex reconstruction efficiency as a function of the jet  $p_T$  (left) and  $\eta$  (right) for b (blue line), c (green line) and light-flavour (red line) jets in  $t\bar{t}$  events [79].

MV2c20 as example, the background sample is composed of 20% c-jets and 80% light-flavour jets. In the MV2c00 the training is performed with only light jets as background. In figure 3.11, a comparison between the MV2cXX background rejections is shown. Most of physics analyses are actually more limited by the c-jets rather than the light-flavour jet rejection, therefore it's performed a training more dedicated to the c-jet rejection.

The output of the MV2cXX tagger is a continuous value; it is possible to choose a threshold value to tag a jet as a b-jet if the MV2cXX output value, for a given jet, is greater than the chosen threshold. The distribution of the MV2c10 discriminant is showed in figure 3.12.

### 3.5 Missing energy reconstruction

Protons in the LHC collisions have a low momenta in the transverse plane respect to the beam axis. So in this plane the momentum conservation is expected, whose value is zero. Missing energy is defined as the energy not detected in a particle detector, but it's needed to balance the momentum conservation.

Missing energy is due to particles that do not interact by electromagnetic or strong forces, and thus leave no trace of their passage in the detectors. In the Standard Model these particles are the neutrinos, but missing energy is expected to be a signature of many theories of physics beyond the Standard

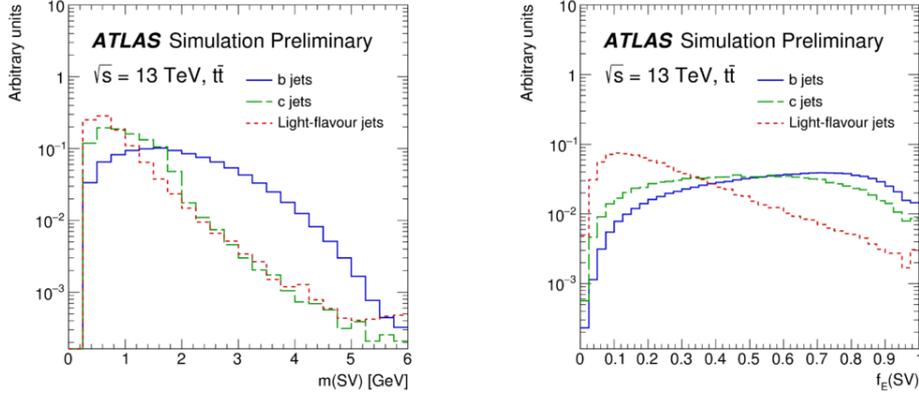


Figure 3.10: Secondary vertex properties reconstructed by the JetFitter algorithm for b (blue line), c (green line) and light-flavour (red line) jets in  $t\bar{t}$  events: Invariant mass of the secondary vertex pair tracks (left) and energy fraction, defined as the ratio of the energy of the secondary vertex pair tracks and energies of all tracks inside the jets (right) [79].

Model.

$E_T^{miss}$  measurements are affected by Standard Model interacting particles that are mis-reconstructed or are beyond the detector acceptance. Several algorithms have been developed for the missing transverse energy reconstruction, like the CST  $E_T^{miss}$ , Track  $E_T^{miss}$  and TST  $E_T^{miss}$  [80, 81].

CST (Calorimeter-based Soft Term) is based on the energy deposit in the ATLAS calorimeter. The measurement includes contributes from hard-objects (reconstructed electrons, photons,  $\tau$  leptons, muons or jets) and also contributes from soft terms related to soft radiation. A weakness of this method is the vulnerability to the pile-up events, that add an additional soft term contribution.

The Track  $E_T^{miss}$  algorithm is based on the track  $p_T$  measurements in the Inner Detector, giving a more robustness with respect to pile-up, but the method is insensitive to neutral particles and the acceptance is limited to the region covered from ID.

The TST method combines informations from calorimeters and tracker, employing track-based soft terms.

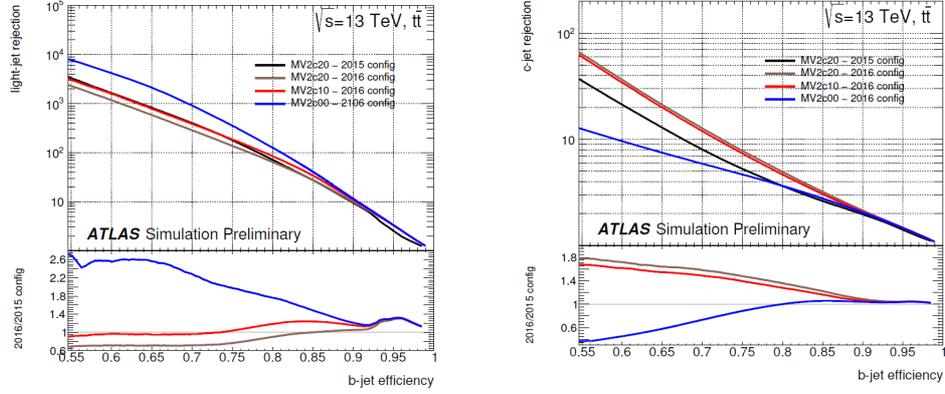


Figure 3.11: Light-flavour jets (left) and c-jets (right) rejections as a function of the b-jets efficiency for the MV2-b algorithms in  $t\bar{t}$  events [79].

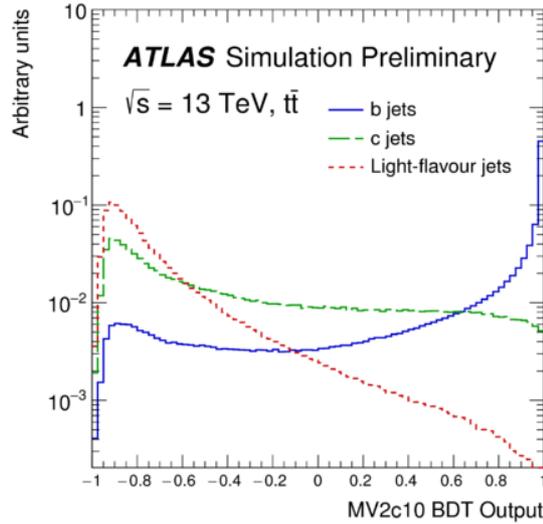


Figure 3.12: Output of the MV2c10 discriminant for b (blue line), c (green line) and light-flavour (red line) jets in  $t\bar{t}$  events [79].

In the Run-I the CST algorithm has been used, while in the Run-II the TST is the method adopted. The transverse missing energy is computed, event by event, as the negative sum of the momenta of all final state objects in the transverse plane:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,soft} \quad (3.6)$$

The terms for the jets, charged leptons and photons are the negative sum of the momenta of the respective calibrated objects. The soft terms could be reconstructed by calorimeter soft term (CST) or track soft term (TST), and the choice affects the performance and the uncertainties in the missing energy reconstruction.

Given the  $E_{x(y)}^{miss}$  components, the  $E_T^{miss}$  and the azimuthal angle  $\phi^{miss}$  are given by:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad (3.7)$$

$$\phi^{miss} = \arctan(E_y^{miss} / E_x^{miss}) \quad (3.8)$$

# Chapter 4

## Search for heavy-diboson resonances in the $llqq$ final state

The search for heavy-diboson resonances in the semileptonic analyses are characterized by a final state with two vector bosons, in which one boson decays leptonically and the other hadronically. Three different final states are possible, depending on the leptonic decay mode of the vector boson:  $X \rightarrow ZV \rightarrow llqq$ ,  $X \rightarrow ZV \rightarrow \nu\nu qq$  and  $X \rightarrow WV \rightarrow l\nu qq$ . The analysis combining all these three channels is going to be published in spring/summer 2020. Previous publications with data collected in the first part of the Run-II (2015 + 2016), corresponding to a luminosity of  $36.1 \text{ fb}^{-1}$ , were performed considering the  $l\nu qq$  final state alone [82], and combining  $llqq$  and  $\nu\nu qq$  final states [83].

My thesis work has been focused on a deep study and on the optimization of the selection for searching heavy-diboson resonances in the semileptonic  $llqq$  final state. In this chapter this search will be exhaustively described. In figure 4.1 a schematic view of the decays is showed. The dataset used in this analysis is the full available Run-2 statistics collected by ATLAS at  $\sqrt{s} = 13 \text{ TeV}$ , corresponding to an integrated luminosity of  $139 \text{ fb}^{-1}$ . Heavy resonances would manifest themselves as resonant structures above the SM background in the invariant-mass distribution of the final state. Upper limits on the production cross sections of heavy resonances times their decay branching ratios to the diboson pair are derived in the mass range 300-5000 GeV, within the context of Beyond Standard Model theories with an heavy vector triplet, or a Gravitons and a Radions in extra dimensions. Production through gluon-gluon fusion, Drell-Yan or vector-boson fusion are considered, depending on the assumed model. Several improvements in the analysis have been introduced with respect to

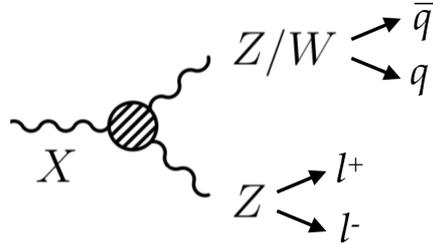


Figure 4.1: Illustration of the diboson semileptonic  $llqq$  final state.

the previous results: the dataset analysed is approximately 4 times larger than the previous one, and several new ideas to improve the sensitivity have been introduced like the use of Track Calo Cluster jets in the boosted regime and the use of Machine Learning techniques in order to categorize the signal events according to the resonance production mechanism. My work has focused mainly on the isolation requirements of the leptons selected and on general improvements of the analysis selection.

Moreover, the  $llqq$  channel has been combined with the other two lepton channels for the final results (thanks to the orthogonality of the corresponding event selections).

## 4.1 Data and Monte Carlo samples

### Data samples

The data sample used has been recorded by the ATLAS experiment between 2015 and 2018 and it corresponds to an integrated luminosity of  $139 \text{ fb}^{-1}$ . The yearly integrated luminosities are summarized in table 4.1. The uncertainty in the combined 2015-2018 integrated luminosity is 1.7% [84], obtained using the LUCID-2 detector [85] for the primary luminosity measurements.

Only datasets satisfying some quality criteria are used: The GoodRunList (GRL) is a list containing all the good quality runs, in which the events have been selected when all ATLAS sub-detectors were fully operating, thus ensuring good data quality.

### Monte Carlo samples

Monte Carlo simulated events have been generated in order to reproduce  $pp$  collisions at a centre of mass energy of  $\sqrt{s} = 13 \text{ TeV}$  and a bunch cross spac-

Year	$\mathcal{L}$ [fb <sup>-1</sup> ]
2015	3.21
2016	32.88
2017	44.31
2018	58.45
total	139.0 ± 2.4

Table 4.1: Breakdown of the integrated luminosity per year.

ing of 25 ns. Monte Carlo samples are used for the background modelling, evaluation of the signal acceptance, optimisation of event selection, estimation of systematic uncertainties and the statistical analysis.

Generated signal and background events are processed through the full ATLAS detector simulation program [86] based on GEANT4 [87]. Additional  $pp$  collisions generated with PYTHIA 8.186 [88] are overlaid to model the effect of the pileup for all MC events. The EvtGen v1.2.0 program [89] is used for the properties of the bottom and charm hadron decays. All simulated events are processed with the same trigger and reconstruction algorithm as the data. Different MC generators have been used in order to simulate signal and background processes. Background processes include the production of a  $W$  or  $Z$  boson in association with jets (denoted as  $W/Z$ +jets), top quark production ( $t\bar{t}$  and single top production) and non-resonant diboson production mainly from the SM  $ZZ$  and  $ZW$ .

Background events containing a  $W$  or a  $Z$  boson with associated jets are simulated using the Sherpa 2.2.1 generator [90]. Matrix elements are calculated for up to 2 partons at next-to-leading order (NLO) and 4 partons at leading-order (LO) using the Comix [91] and OpenLoops [92] matrix element generators, merged with the Sherpa parton shower [93] using the ME+PS@NLO prescription [94]. The  $W/Z$  + jets MC predictions are generated using NNPDF3.0NLO [95] parton distribution functions sets and are calculated to the NNLO cross sections [96]. The diboson processes are also generated with Sherpa 2.2.1 generator [90].

The  $t\bar{t}$  and single-top events are generated with NLO accuracy using the Powheg-Box [97] generator with the NNPDF3.0NLO [95] parton distribution functions sets in the matrix element calculation. The top quark spin correlations are preserved (for t-channel, top quarks are decayed using MadSpin [98]). For all processes the parton shower, fragmentation, and the un-

derlying events are simulated using PYTHIA8.230 with the A14 tune set [99]. The top mass is set to 172.5 GeV. The cross sections of  $t\bar{t}$  and single-top, known to NNLO in QCD including re-summation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms[100, 101, 102, 103], are used to normalise the simulated samples.

Signal samples for the gluon-gluon-fusion RS Graviton and HVT (both Drell-Yan and Vector Boson Fusion production mechanisms) are generated with MADGRAPH5-2.2.2 (MG5) [104] interfaced to PYTHIA 8.186. The Radion signals (both ggF and VBF) and VBF RS Graviton samples are instead produced with MADGRAPH5-2.6.0 (MG5) interfaced to PYTHIA 8.212.

The RS Graviton samples are generated with  $k/\overline{M}_{Planck} = 1$  and the Radion samples are generated with  $k\pi r_c = 35$  and  $\Lambda_R = 3$  TeV. The HVT model A signal samples are generated with  $g_H = -0.56$  and  $g_F = -0.55$ , while the model B signal samples with  $g_H = -2.9$  and  $g_F = 0.14$ . In model B, the  $V'$  resonances are broader than the ones in the weakly coupled scenario, model A, but remain narrow relative to the experimental resolution, so the model B interpretation is performed assuming the same signal shape as the model A.

Another set of HVT signal samples is generated in the model C with  $g_H = 1$  and  $g_F = 0$  for generation of resonances produced via VBF.

Masses of diboson resonances are varied from 300 GeV to 5 TeV for each scenario.

## 4.2 Object reconstruction

The final state of the  $llqq$  analysis consists of two vector bosons, where one decays hadronically and the other leptonically. Lepton (i.e electrons and muons) and jet identification and reconstruction are key components of the  $llqq$  analysis. The selection requirements applied are reported in this section.

### 4.2.1 Leptons

‘Loose’ electrons and muons are used to select  $Z \rightarrow e^+e^- / Z \rightarrow \mu^+\mu^-$  candidates. The selections for the leptons candidates are summarized in table 4.2.

Leptons are required to have associated tracks satisfying  $|d_0/\sigma_{d_0}| < 5$  (3) for electrons (muons) and  $|z_0 \times \sin \theta| < 0.5$  mm both for electrons and muons. The parameter  $d_0$  is the transverse impact parameter with respect to the beam line and  $\sigma_{d_0}$  is its uncertainty;  $z_0$  is the distance from the primary

	Electrons	Muons
Lepton ID	<i>Loose</i>	<i>Loose</i>
Pseudorapidity range	$ \eta  < 2.47$	$ \eta  < 2.5$
Transverse momentum	$p_T > 7$ GeV	$p_T > 7$
Track to vertex association	$ d_0/\sigma_{d_0}  < 5$	$ d_0/\sigma_{d_0}  < 3$
	$ z_0 \times \sin \theta  < 0.5$ mm	$ z_0 \times \sin \theta  < 0.5$ mm
Isolation	FixedCutLoose at $p_T < 100$ GeV and no isolation requirement at $> 100$ GeV	FixedCutLoose at $p_T < 100$ GeV and no isolation requirement at $> 100$ GeV

Table 4.2: Summary of leptons selections.

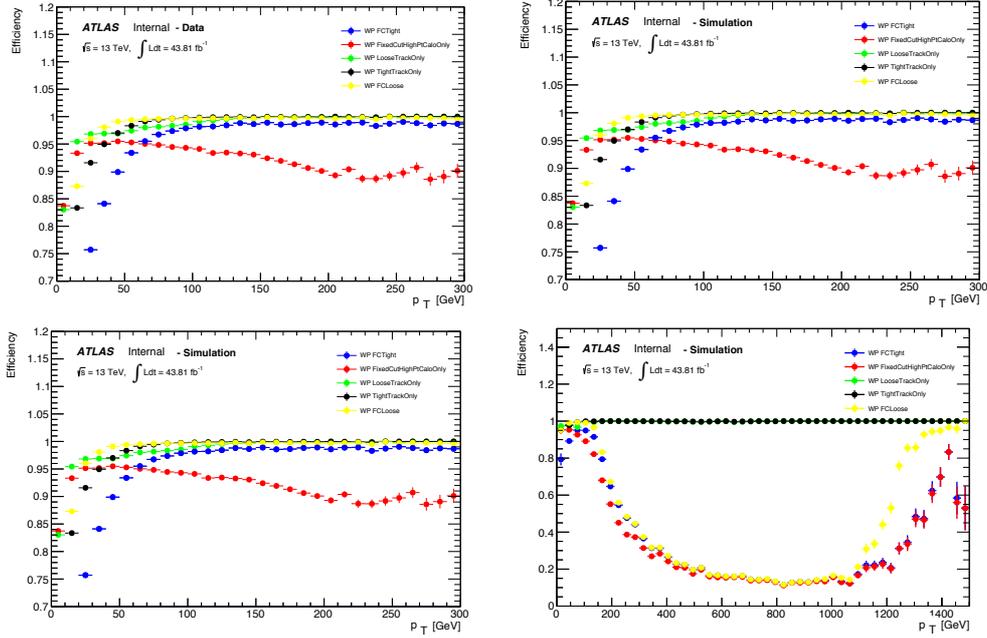


Figure 4.2: Lepton identification efficiencies in the electron channel, testing various isolation WPs, as a function of the probe lepton  $p_T$  using data 2017 (upper left),  $Z$ +jets (upper right),  $ggF H \rightarrow ZZllqq$  with  $m = 1$  TeV (lower left) and  $H \rightarrow ZZllqq$  with  $m = 3$  TeV (lower right) samples.

vertex of the longitudinal position of the track along the beam line at the point where  $d_0$  is measured.

Leptons from  $W$  and  $Z$  boson decays are expected to be isolated from other energy deposits in the detector. Thus isolation criteria based on the sum of track  $p_T$ , the sum of calorimeter  $E_T$ , or both in small cones around the lepton direction are used to further reduce backgrounds from non-isolated sources. Looser isolation working points based on the ID track variables only are preferable to keep high signal efficiency in the high- $p_T$  region, where two

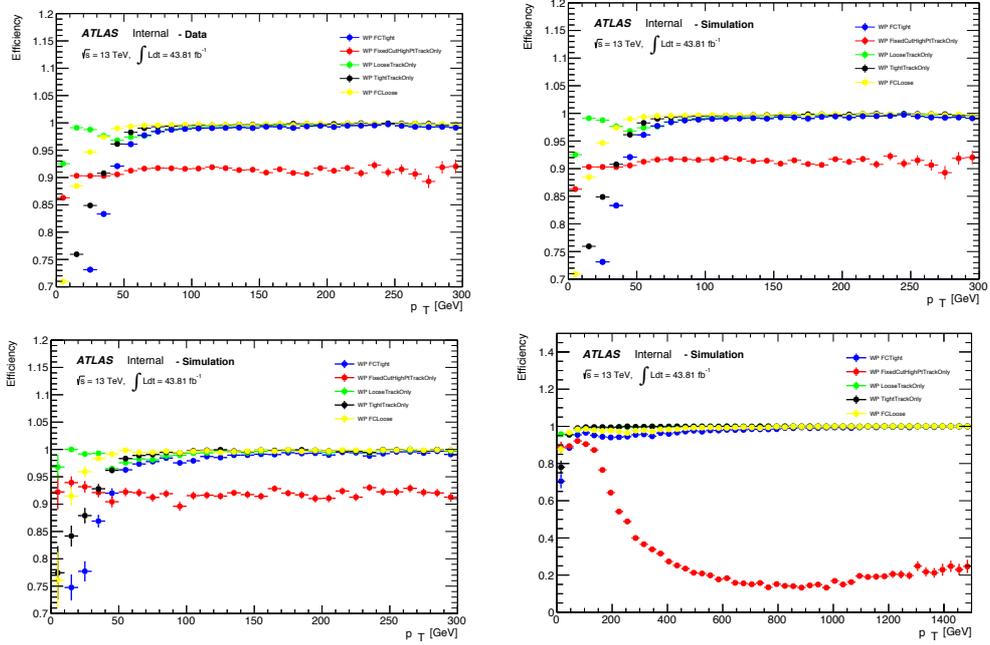


Figure 4.3: Lepton identification efficiencies in the muon channel, testing various isolation WPs, as a function of the probe lepton  $p_T$  using data 2017 (upper left),  $Z$ +jets (upper right),  $ggF H \rightarrow ZZllqq$  with  $m = 1$  TeV (lower left) and  $H \rightarrow ZZllqq$  with  $m = 3$  TeV (lower right) samples.

lepton candidates are close to each other and their calorimetric energy deposits happen to overlap.

In the previous analysis the `LooseTrackOnly` working point (see sections 3.1 and 3.2) was used for lepton isolation, but for the current full Run-2 analysis this working point is not supported anymore. Part of my work has been a detailed study in order to select the best isolation condition.

I measured the efficiencies of some WPs available for the lepton isolation requirements, by using the tag-and-probe method on  $Z \rightarrow \ell\ell$  candidates in the data collected in 2017 as well as simulated events for  $Z$ +jets and signal samples. The selections required to identify the  $Z \rightarrow \ell\ell$  candidates in this study are summarized in table 4.3.

Same-flavor dilepton candidates with the invariant mass consistent with  $Z$ -boson mass have been selected. If more than one dilepton candidate exists, the one with the invariant mass closer to the  $Z$  mass is selected. To avoid the bias of the selections applied in the single-lepton trigger algorithm, one of the two lepton candidates (tag) is required to be ‘Tight’, to pass the

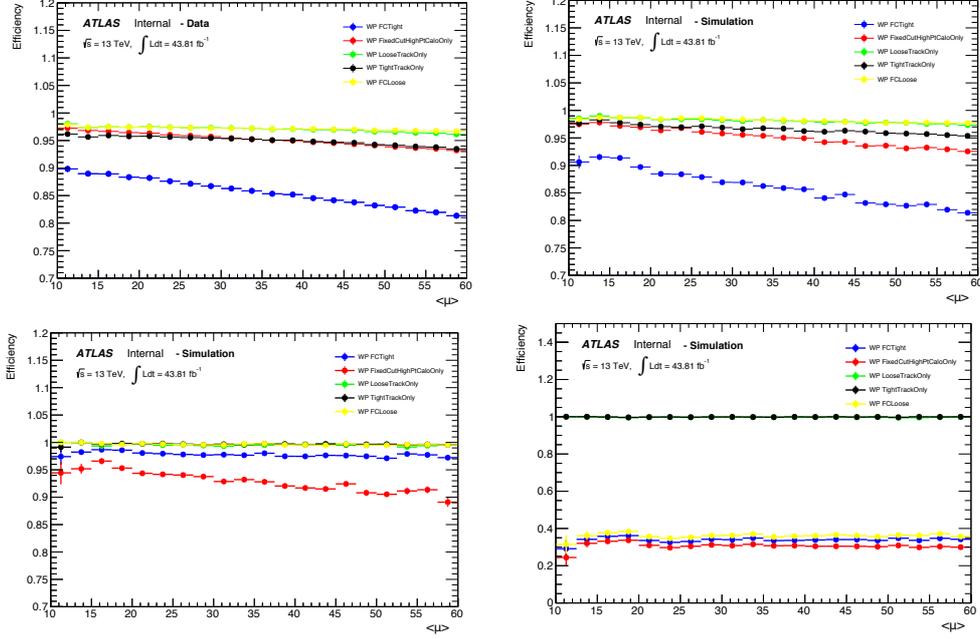


Figure 4.4: Lepton identification efficiencies in the electron channel, testing various isolation WPs, as a function of the pile-up for data 2017 (upper left),  $Z$ +jets (upper right),  $ggF H \rightarrow ZZllqq$  with  $m = 1$  TeV (lower left) and  $H \rightarrow ZZllqq$  with  $m = 3$  TeV (lower right) samples.

`FixedCutTightTrackOnly` isolation, to have  $p_T > 28$  GeV and match to the lepton used to trigger the event. Opposite lepton (probe) is used to measure the isolation efficiency. Figures 4.2, 4.3, 4.4, 4.5 show the efficiencies for the various WPs tested in the electron and muon channels, as a function of the probe lepton  $p_T$  and the pile-up  $\langle \mu \rangle$ , respectively, using the data collected in 2017, simulated  $Z$ +jets and  $ggF$  NWA Higgs with resonance masses of 1 TeV and 3 TeV.

For 3 TeV signal, the isolation efficiencies using the calo-cluster informations are extremely low, especially in the electron channel. This could be due to the smaller angular separation between the electrons. Figure 4.6 shows the angular  $\Delta R$  separation between the two electrons for signal  $ggF$  NWA Higgs with  $m = 1$  TeV and 3 TeV. At  $m = 3$  TeV, due to the increased mass value, the electrons are more collimated and, therefore, not isolated. Also the efficiencies as a function of lepton  $p_T$  for 3 TeV signal, after requiring a dilepton separation  $\Delta R > 0.3$ , are showed. In this condition the electrons result well isolated between each other also for the calorimetric working points.

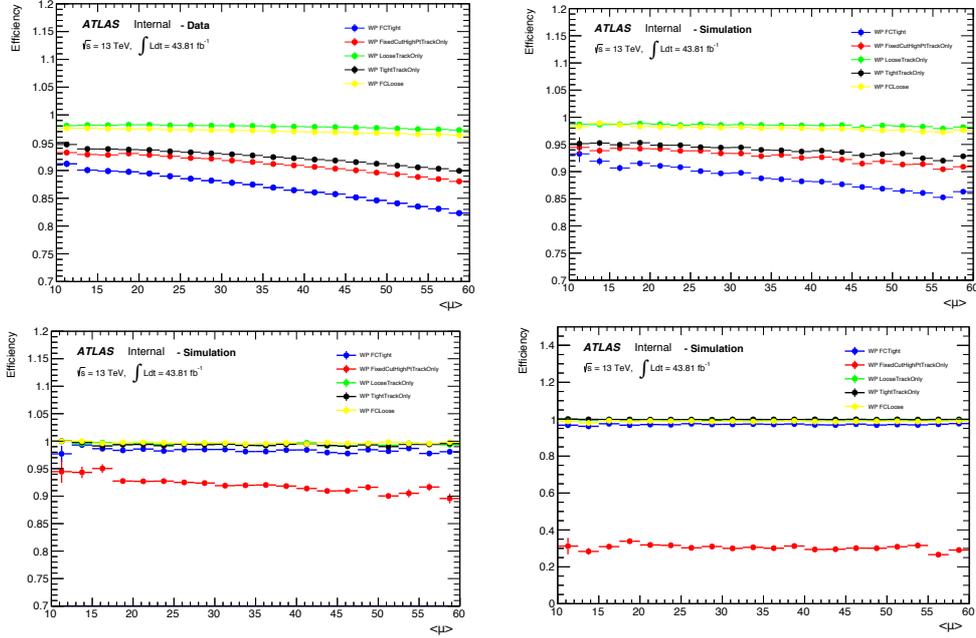


Figure 4.5: Lepton identification efficiencies in the muon channel, testing various isolation WPs, as a function of the pile-up for data 2017 (upper left),  $Z$ +jets (upper right),  $ggF H \rightarrow ZZllqq$  with  $m = 1$  TeV (lower left) and  $H \rightarrow ZZllqq$  with  $m = 3$  TeV (lower right) samples.

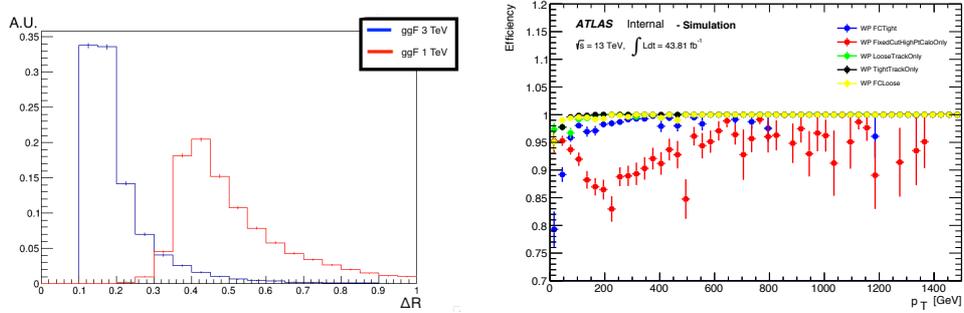


Figure 4.6:  $\Delta R$  between the electrons of the dilepton couple for signal  $ggF$  1 TeV, in red, and 3 TeV, in blue (plot on the left). Lepton identification efficiencies in the electron channel, testing various isolation WPs, as a function of the probe lepton  $p_T$  for  $ggF H \rightarrow ZZllqq$  3 TeV samples, requiring a dilepton separation  $\Delta R > 0.3$  (plot on the right).

The change in the number of background events and the number of sig-

N Loose Lepton	$> 1$
$\eta$ leptons	$\leq 2.5$
Opposite sign of muon channel	
Invariant Mass	$83 < M_{ee} < 99$ or $85.6 - 0.0117p_T^{\mu\mu} < M_{\mu\mu} < 94.0 + 0.0185p_T^{\mu\mu}$
Tag Lepton	FixedCutTightTrackOnlyIso and TightLH and Lepton $p_T > 28$ GeV

Table 4.3: Event selection for tag and probe analysis. Triggers are the same of the analysis selection and they will be discussed in section 4.3.1.

nal events (testing different signal hypotheses), applying different isolation requirements, has been investigated both after the identification of the leptonic  $Z$ -decay and in the final regions defined in the analysis selection (see section 4.3). The background events are not increased even if no isolation cuts are applied in the region where the lepton  $p_T$  is greater than 100 GeV. In order to keep an high signal efficiency in the high- $p_T$  region (where the leptons are more collimated) and reduce the background contamination in the low lepton  $p_T$  region, the `FixedCutLoose` isolation condition is applied only at  $p_T < 100$  GeV.

## 4.2.2 Jets

In this analysis both small-R jets and large-R jets are used. As the momentum of the boson increases, the quarks pair is increasingly boosted, so if the resonance mass is significantly higher than the  $V$  boson mass, the  $qq$  pair from the  $V$  boson decay can be collimated. In this case hadrons from the two quarks overlap in the detector and are more efficiently reconstructed as a single large-R jet. In the semileptonic analyses a mass range from 300 to 5000 GeV is investigated, and two different reconstruction techniques for the  $V \rightarrow qq$  decay are considered: resolved and merged. The resolved reconstruction attempts to identify two separate small-R jets of hadrons from the  $V$  decay, while merged reconstruction identifies the  $V \rightarrow qq$  decay as a single large-R jet, using jet substructure techniques.

### Small-R Jets

Small-R jets are used not only to reconstruct the  $W/Z \rightarrow qq$  candidates that are less boosted, so that the  $qq$  pair from the  $V$  boson decay are well separated, but also to select the forward jets coming from the vector-boson fusion production mechanism. They are reconstructed from calorimeter clus-

ters using the *anti* -  $k_T$  algorithm with radius parameter  $R = 0.4$ . Jets are required to have  $p_T > 30$  GeV and  $|\eta| < 4.5$ . To suppress jets from pile-up interactions, a jet vertex tagger is applied to jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$ , based on informations about tracks associated with the primary vertex and pile-up vertices. Small-R jets containing b-hadrons are identified using the multivariate *MV2c10* algorithm, and the identification is limited to  $|\eta| < 2.5$  due to the ID coverage. Small-R jet selection criteria are summarized in table 4.4.

Jet reconstruction parameters		
Parameter	Value	
algorithm	anti- $k_T$	
R-parameter	0.4	
input constituent	EMTopo	
Selection requirements		
	central jet	forward jet
Observable	Requirement	
$p_T$	$>30$ GeV	
$ \eta $	$<2.5$	$< 4.5$
JVT	$> 0.59$ for $p_T < 120$ GeV and $ \eta  < 2.4$ $> 0.11$ for $p_T < 120$ GeV and $2.4 <  \eta  < 2.5$ (Medium working point)	
<i>b</i> -tagging	Tagged, or not tagged	Not tagged

Table 4.4: Summary of small-R jet selection criteria.

### Large-R Jets

Large-R jets are reconstructed from the Track-Calo Clusters (TCCs) using the *anti* -  $k_T$  algorithm with radius parameter  $R = 1.0$ . As discussed in section 3.3, TCC algorithm employs the spatial coordinates of the tracker and the energy scale of the calorimeter, and this results in improved resolution for both jet mass and substructure variables. The trimming procedure is applied to suppress the pileup and soft radiations contributions. Large-R jets are required to have  $p_T > 200$  GeV,  $|\eta| < 2.0$  and a jet invariant mass  $m_J > 50$  GeV. The reconstructed  $m_J$  peaks around the  $m_{W/Z}$  for  $W/Z \rightarrow q\bar{q}$  signals, while distributes broadly for single-quark jets and gluon-induced jets. Discrimination against background jets is performed by defining a mass

window around the  $W/Z$  mass and by cutting on the variable  $D_2$ , which is defined as a ratio of two-point to three-point energy correlation functions, that are based on the energies and pairwise angular distances between the jet constituents [105, 106]:

Jet reconstruction parameters	
Parameter	Value
algorithm	anti- $k_T$
R-parameter	1.0
input constituent	TrackCaloCluster
grooming algorithm	Trimming
$f_{cut}$	0.05
$R_{trim}$	0.2
Selection requirements	
Observable	Requirement
$p_T$	>200 GeV
$ \eta $	<2.0
mass	> 50 GeV

Table 4.5: Summary of large-R jet selection criteria.

$$D_2^{(\beta=1)} = E_{CF3} \left( \frac{E_{CF1}}{E_{CF2}} \right)^3 \quad (4.1)$$

where the energy correlation functions ( $E_{CF}$ ) are defined as:

$$E_{CF1} = \sum_i p_{T,i} \quad (4.2)$$

$$E_{CF2} = \sum_{ij} p_{T,i} p_{T,j} \Delta R_{ij} \quad (4.3)$$

$$E_{CF3} = \sum_{ijk} p_{T,i} p_{T,j} p_{T,k} \Delta R_{ij} \Delta R_{jk} \Delta R_{ki} \quad (4.4)$$

A two-dimensional optimization of the jet mass windows and  $D_2$  thresholds is performed in order to enhance the signal to background separation and maximize the analysis sensitivity. Figure 4.7 shows the optimized thresholds on  $D_2$  and jet mass, as a function of the jet  $p_T$ . The  $W/Z$  taggers are optimized to achieve about 40-50% efficiency at the lowest- $p_T$  region, about 60% at the intermediate  $p_T$  range and about 70% at the highest- $p_T$  region. Signal efficiency and background rejection, compared with the LCTopo jet, are

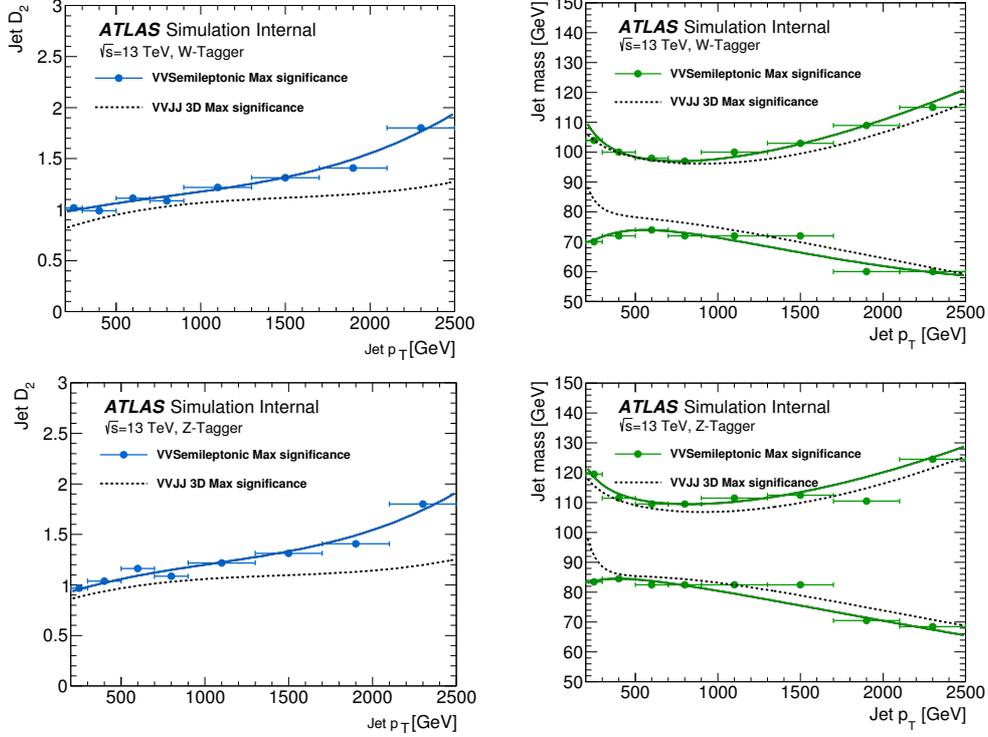


Figure 4.7: The upper cut on  $D_2$  (left plots) and jet mass window (right plots) of the  $W$ -tagger (upper plots) and  $Z$ -tagger (lower plots) as a function of the large- $R$  jet  $p_T$ . The optimal cut values for maximum significance are shown as solid markers and the fitted function as solid lines. Working points from  $VV \rightarrow JJ$  analysis [107] are also shown as dashed lines as a reference.

shown in Figure 4.8. TCC jet tagger achieved higher signal efficiency while keeping similar background rejection to LCTopo jet tagger at highest- $p_T$  region.

Figure 4.9 shows the comparison of TCC jet mass and  $D_2$  resolutions with LCTopo jets, as a function of  $p_T$ .

The large- $R$  jet selection criteria are summarized in table 4.5.

### Variable Radius (VR) Track jets

The variable-radius (VR) track jets [108] are used to identify large- $R$  jets containing  $b$ -hadrons. VR jets are reconstructed using the anti- $k_T$  algorithm

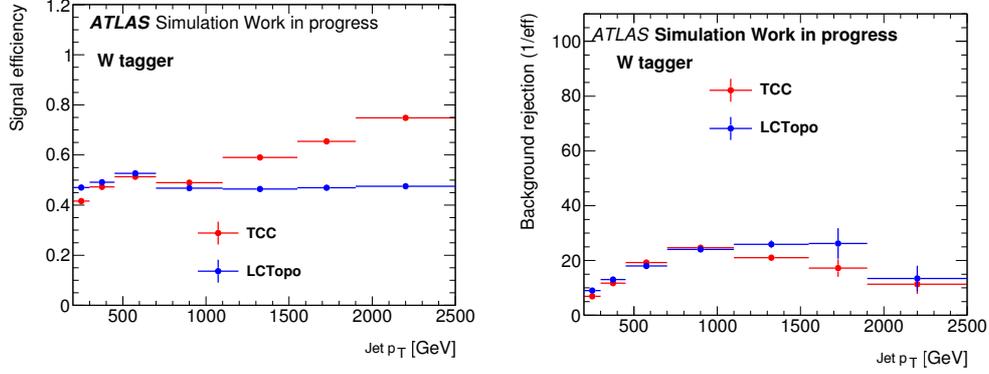


Figure 4.8: Signal efficiency (left plot) and background rejection (right plot) comparison between TCC jets and LCTopo jets.

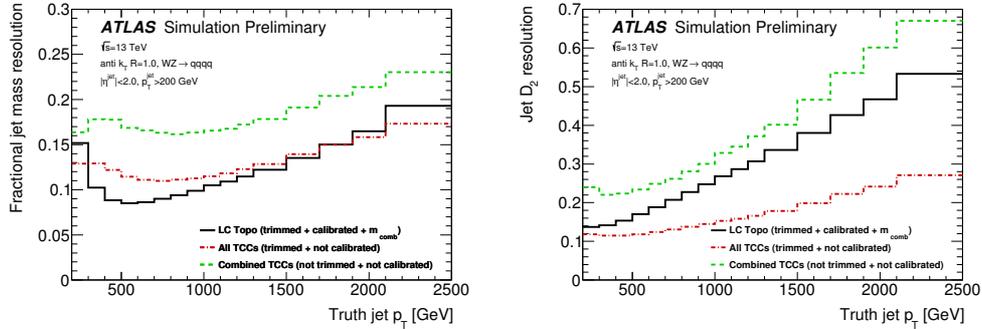


Figure 4.9: A comparison of the fractional jet mass (left plot) and  $D_2$  (right plot) resolution for topo-cluster jets (solid black lines), jets built using all-TCC objects (dash-dotted red lines) and jets built using only combined-TCC objects (dashed green lines), as a function of truth-jet  $p_T$  [69]. The TCC definitions are seen to provide the best resolution at high- $p_T$ , particularly for  $D_2$ .

from the ID tracks associated with large-R jets. Their radius varies with track jet  $p_T$  as:

$$R_{\text{eff}}(p_T) = \frac{\rho}{p_T}. \quad (4.5)$$

The parameter  $\rho$  is fixed to 30 GeV and upper and lower limit of the cone size are set to  $R_{\text{max}} = 0.4$  and  $R_{\text{min}} = 0.02$ . The VR jets are required to satisfy  $p_T > 10$  GeV and  $|\eta| < 2.5$ .

### 4.2.3 Overlap Removal

A particle detected in ATLAS can be reconstructed by several reconstruction algorithms. As example the same energy deposit can be reconstructed both as an electron and a jet. In order to avoid double-counting, an overlap-removal procedure is adopted to the selected leptons and jets. The following criteria are applied:

- For nearby electrons and small-R jets: the jet is removed if the separation between the electron and jet is within  $\Delta R < 0.2$ , otherwise the electron is removed if the separation is within  $0.2 < \Delta R < 0.4$ .
- For nearby muons and small-R jets: the jet is removed if the separation between the muon and jet is within  $\Delta R < 0.2$  and if the the jet has less than 3 tracks or the energy and momentum differences between the muon and the jet are small. The muon is otherwise removed if the separation is within  $\Delta R < 0.4$ .
- For nearby electrons and large-R jets: the large-R jet is removed if the separation between the electron and the large-R jet is within  $\Delta R < 0.1$ .
- For nearby electrons sharing the same ID track: the electron with the smallest  $p_T$  is removed.
- For nearby electrons and muons sharing the same ID track: if the muon is a *calo-muon* the electron is removed, otherwise the muon is removed.

## 4.3 The event selection

Multiple signal regions (SRs) are defined in order to maximize the expected significance of the signal over the background.

The events selection starts with the identification of the  $Z \rightarrow ll$  decay. The

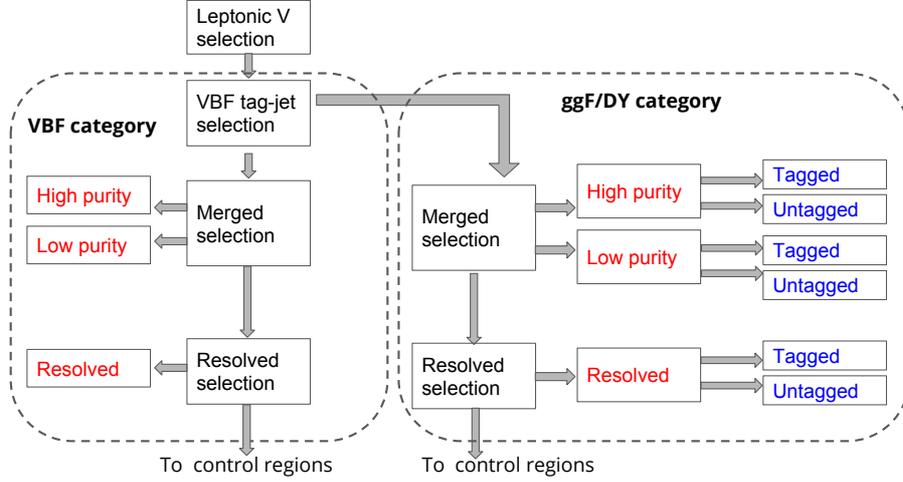


Figure 4.10: Illustration of the analysis flow.

events are then categorized according to the resonance production mechanism: two categories are defined, the VBF category and the ggF/DY category. For each of these categories the selection proceeds with the identification of the V-hadronic decay. The merged regime is investigated first, looking for the  $V \rightarrow J$  decay and, if it fails, the analysis flow looks for the resolved  $V \rightarrow qq$  decay. A further categorization is applied in the merged regime according to the  $W/Z$  tagger working points, defining *High Purity* and *Low Purity* regions.

Depending on the spin hypothesis, the analysis looks for  $ZZ \rightarrow llqq$  and  $ZW \rightarrow llqq$  candidates. About 21% of the  $Z$  – hadronic decays have two b-quarks, whereas the dominant background,  $Z + jets$ , has a smaller heavy-quark content. For this reason, in the spin-0/spin-2 interpretations and for the ggF/DY category, two different regions are further defined according to the number of b-hadrons: *Tagged* and *Untagged* regions.

Since no enhancement of b-quark content is expected from the  $W$  – hadronic decay, the split is not done in the case of spin-1 interpretation.

Also no split is done, both in the  $ZZ/ZW \rightarrow llqq$  selections, for the VBF category due to low statistics.

Events not satisfying the SRs selections are used in the Control Regions (CRs) for the background estimation. Figure 4.10 shows a schematic illustration of the selection flow.

The different steps of the analysis selection will be described in the following

sections.

### 4.3.1 Trigger requirements

The data used in this analysis have been collected using a combination of multiple single-electron (for the search in the  $eeqq$  final state) and single-muon triggers (for the search in the  $\mu\mu qq$  final state), with varying transverse momentum thresholds, quality and isolation requirements.

Data-taking period	Single Electron triggers	Single Muon Triggers
2015	HLT e24 lhmedium L1EM20 HLT e60 lhmedium HLT e120 lhloose	HLT mu20 iloose L1MU15 HLT mu50
2016	HLT e26 lhtight nod0 ivarloose HLT e60 lhmedium nod0 HLT e140 lhloose nod0	HLT mu26 ivarmedium HLT mu50
2017	same as above	same as above
2018	same as above	same as above

Table 4.6: List of the trigger used in the  $llqq$  analysis. HLT stands for High Level Trigger, then the threshold (i.e. e20), quality (i.e. lhtight) and isolation (i.e. ivarloose) are reported. The nod0 string indicates that no transverse impact parameter cuts are required.

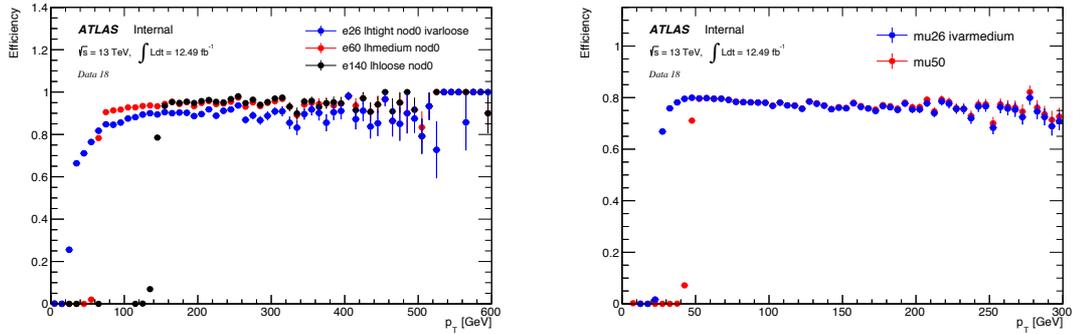


Figure 4.11: Efficiency as a function on the  $p_T$  for electrons (on the left) and muons (on the right) for different trigger thresholds, using the data collected in 2018.

The trigger thresholds of the single-electron triggers varied from 24 GeV to 140 GeV in  $E_T$ , and from 20 GeV to 50 GeV in  $p_T$  for single-muon triggers.

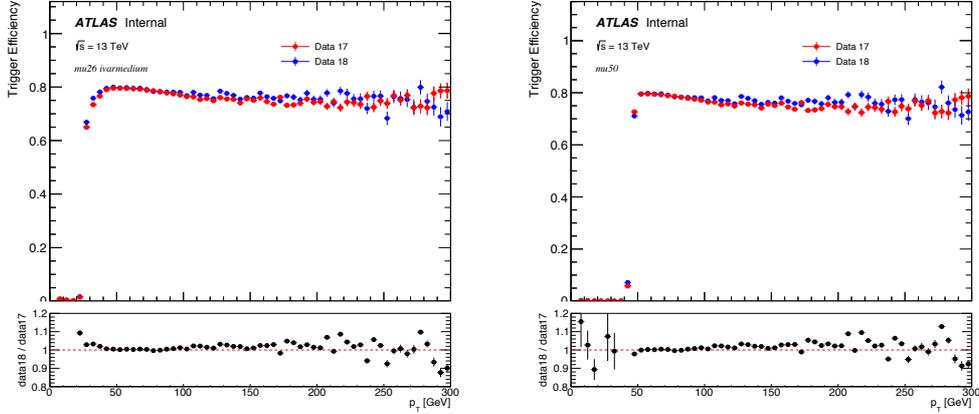


Figure 4.12: Efficiency comparison between the data collected in 2017 and 2018, as a function on the muon  $p_T$ , for the different trigger thresholds.

Triggers with low thresholds have tighter lepton isolation requirements, while triggers with high thresholds have looser or no isolation requirements. The list of the trigger used is reported in the table 4.6.

In figure 4.11 the efficiencies for the different trigger thresholds are shown as a function of the lepton  $p_T$ , using the data collected in 2018, while in figure 4.12 and 4.13 a comparison of the efficiencies for the data collected in 2017 and 2018 is reported.

The single electron triggers have a plateau efficiency of about 90%, while the single muon trigger of about 80%. In the di-muon final state the efficiencies are greater than 90% since there are two muon candidates in the event ( $\epsilon = 1 - (1 - 0.8)^2 > 0.9$ ). Dedicated studies based on simulated data show that the efficiency to trigger signal events is greater than 90% for the entire mass range investigated in this analysis.

The efficiencies have been evaluated using the tag-and-probe method. Events with two same flavour leptons and with the dilepton invariant mass  $m_{\ell\ell}$  consistent with the  $Z$  boson mass are selected. The tag lepton is required to have a  $p_T > 30$  GeV.

### 4.3.2 $Z \rightarrow ll$ selection

The analysis selection starts with the identification of the events containing the leptonic decay of the  $Z$  boson. The  $Z \rightarrow ll$  decays are identified by requiring exactly two isolated same-flavour leptons, satisfying the ‘loose’ cri-

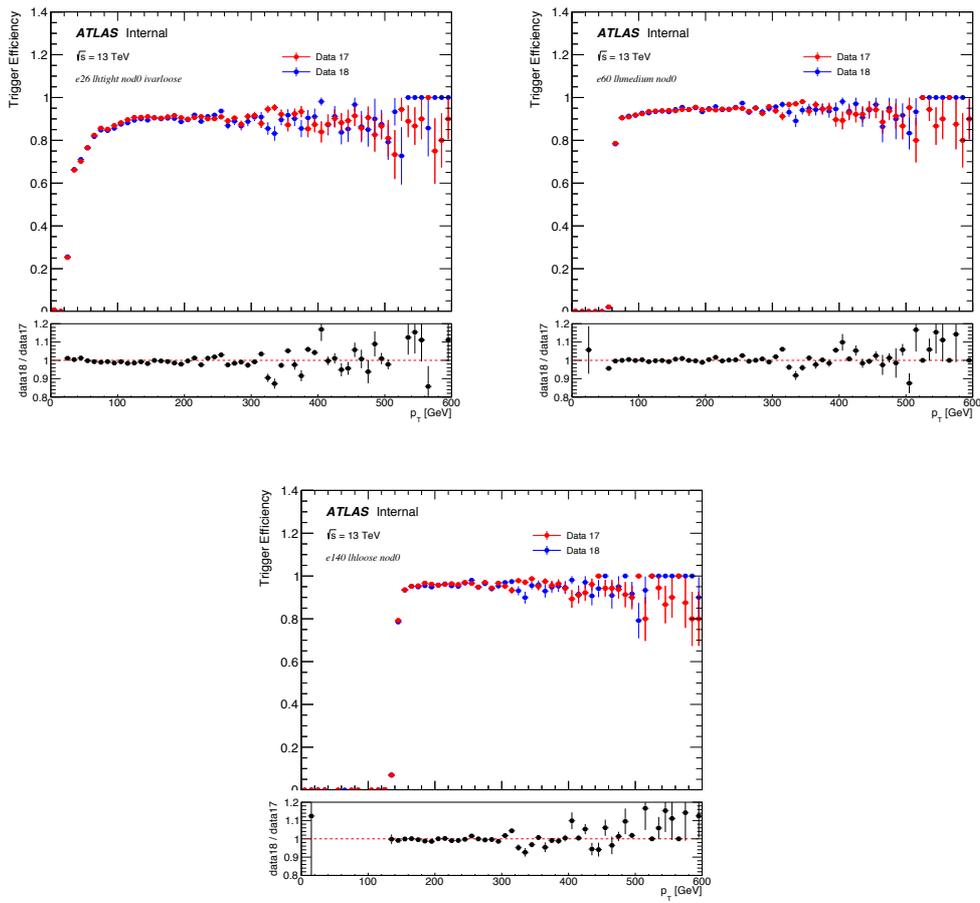


Figure 4.13: Efficiency comparison between the data collected in 2017 and 2018, as a function on the electron  $p_T$ , for the different trigger thresholds.

teria, and vetoing events with additional leptons passing the ‘loose’ criteria. Only electrons and muons are considered, since the  $\tau$  leptons quickly decay and do not have a signature as clear as the other two leptons. Opposite charges are required for the muon pairs but not for the electron pairs, since electrons are more susceptible to charge misidentification due to the conversions of photons from bremsstrahlung, especially at high  $E_T$ .

I performed a detailed study of the  $p_T$  thresholds to be applied for the two selected leptons. Both the two leptons are required to have a  $p_T$  greater than 30 GeV. In the previous analysis [83] the applied cuts on the leading and subleading leptons  $p_T$  were  $p_T^{lead} > 28$  GeV and  $p_T^{sublead} > 7$  GeV. A fake lepton contamination is estimated to be very small, about 0.2% and 0.03% of the total background in di-electron and di-muon channels, respectively. However, as shown in Figure 4.14, the distribution of the subleading lepton  $p_T$  in  $W + jets$  sample (dominated by fake leptons) is significantly lower than that in  $Z + jets$  sample (enriched in real leptons). The distribution shows that a tighter cut on the subleading lepton  $p_T$  can reduce the fake lepton contamination and improve the search sensitivity. It is found that the significance is improved by 4-7% at higher-mass region by the tighter  $p_T$  cut ( $> 30$  GeV for both leptons), while less than 1% sensitivity loss is found at the lowest-mass region.

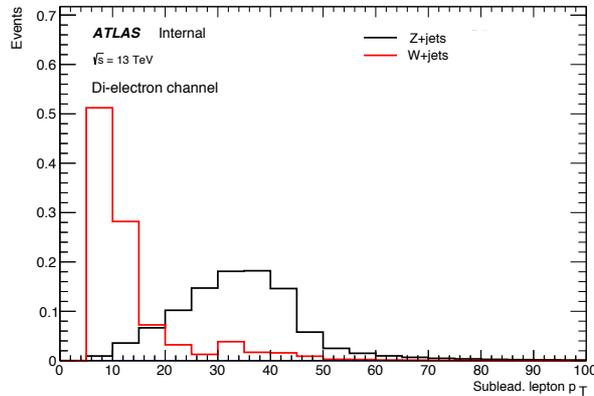


Figure 4.14: Distributions of the subleading lepton  $p_T$  for the  $W + jets$  (in red) and  $Z + jets$  (in black) samples in the di-electron channel. Both distribution are normalized to unity.

The dilepton invariant mass is required to be consistent with the  $Z$  boson mass. For electrons, a fixed window cut is applied:

$$83 \text{ GeV} < m_{ee} < 99 \text{ GeV} \quad (4.6)$$

while for the muons, a  $p_T^{\mu\mu}$ -dependent cut is applied, to take into account the effect of di-muon mass resolution degradation at high transverse momentum:

$$(85.6 - 0.0117 \times p_T^{\mu\mu}) \text{ GeV} < m_{\mu\mu} < (94.0 + 0.0185 \times p_T^{\mu\mu}) \text{ GeV} \quad (4.7)$$

These mass windows are optimized in order to ensure approximately 95% efficiency for the  $Z \rightarrow ll$  selection.

### 4.3.3 The production mechanism categorization

Different resonance production mechanisms are taken into account depending also on the assumed signal model: Gluon–gluon fusion (ggF), Drell–Yan (DY) and vector-boson fusion (VBF) processes are considered. The Feynman diagrams of these models are shown in figure 1.2. Although the production of signal events in VBF mode has lower cross section than the other production mechanisms, the VBF events have more informations since in addition to the presence of a pair of vector bosons from the resonance decay, two additional jets are expected in the forward region of the detector. A good events classification can lead to an increase of signal sensitivity. The features of these additional jets are the large separation in pseudorapidity and the large dijet invariant mass.

In the previous analysis the VBF events were identified requiring two opposite forward jets, satisfying the following criteria:

- Non b-tagged jets;
- $|\eta| < 4.5$  for both jets;
- $\eta_{J_1} \cdot \eta_{J_2} < 0$  (opposite jets);
- $p_T > 30 \text{ GeV}$  for both jets;
- $m_{J_1 J_2} > 700 \text{ GeV}$ ;
- $\Delta\eta_{J_1 J_2} > 4.7$

A bi-dimensional selection in the  $m_{J_1 J_2} - \Delta\eta_{J_1 J_2}$  plane has been performed in order to categorize the events, optimizing the purity of the VBF signal while rejecting the background. The figure of merit taken into account was:  $\epsilon = \epsilon_{VBF} \times (1 - \epsilon_{Bkg}) \times (1 - \epsilon_{ggF})$ , where  $\epsilon_{VBF}$ ,  $\epsilon_{ggF}$ ,  $\epsilon_{Bkg}$  are the selection efficiencies for the VBF signal events, ggF signal events and for the background events, respectively.

With this *cut based* approach the VBF selection efficiency in the VBF category is only 37.8% (27.0%) for spin-0 1 TeV (3 TeV) resonance mass. In order to increase the efficiency to identify VBF events and correctly flag them as belonging to the VBF category, a Machine Learning approach has been implemented in the current analysis. The architecture that has been chosen for the VBF/ggF classification problem is the Recurrent Neural Network (RNN) [109]. The main idea of the NN-based VBF/ggF classification is to use the small-R jets informations present in the event in order to identify the VBF topology, instead of applying selection cuts on a selected pair of jets.

The input features are the 4 components of the 4-momentum of the jets ( $p_T$ ,  $\eta$ ,  $\phi$ ,  $E$ ) that are given as RNN inputs in a recurrent way, according to the number of jets in the events. The maximum number of input jets has been fixed to 2 in order to avoid that same extra systematics on the modelling of an higher numbers of jets could affects the performance and the results of the analysis. Moreover, in order to not consider the jets from the  $W/Z$  decays, the leading and subleading jets in the events are excluded.

The algorithm has been trained on the VBF and ggF spin-0 signal samples with 1 TeV mass values. The model trained on these hypotheses has been used also on others signal mass values, as the VBF topology is expected to be not dependent on the signal mass. The performance for other spin hypotheses have been found good and similar to the spin-0 case, so the same model derived for the spin-0 has been used for all the spin hypotheses.

In figure 4.15 and 4.16 the RNN input variables are shown, comparing the shapes of the VBF and ggF topologies for the two jets used as input to the RNN. The VBF events are all the ones with the RNN score greater to a selected threshold. The value chosen is 0.8, that gives the same background rejection in VBF category as the cut-based approach and harmonize between the other two lepton channels analyses ( $l\nu qq$  and  $\nu\nu qq$ ), that also use this categorization approach. In figure 4.17 are shown the RNN score distributions and the fractions of simulated events passing the RNN requirement as functions of the resonance mass, for different signal models.

The efficiency to select VBF events, with the RNN approach, is found to

be 42.7% and 44.2% for spin0 1 TeV and 3 TeV signal mass hypotheses, respectively.

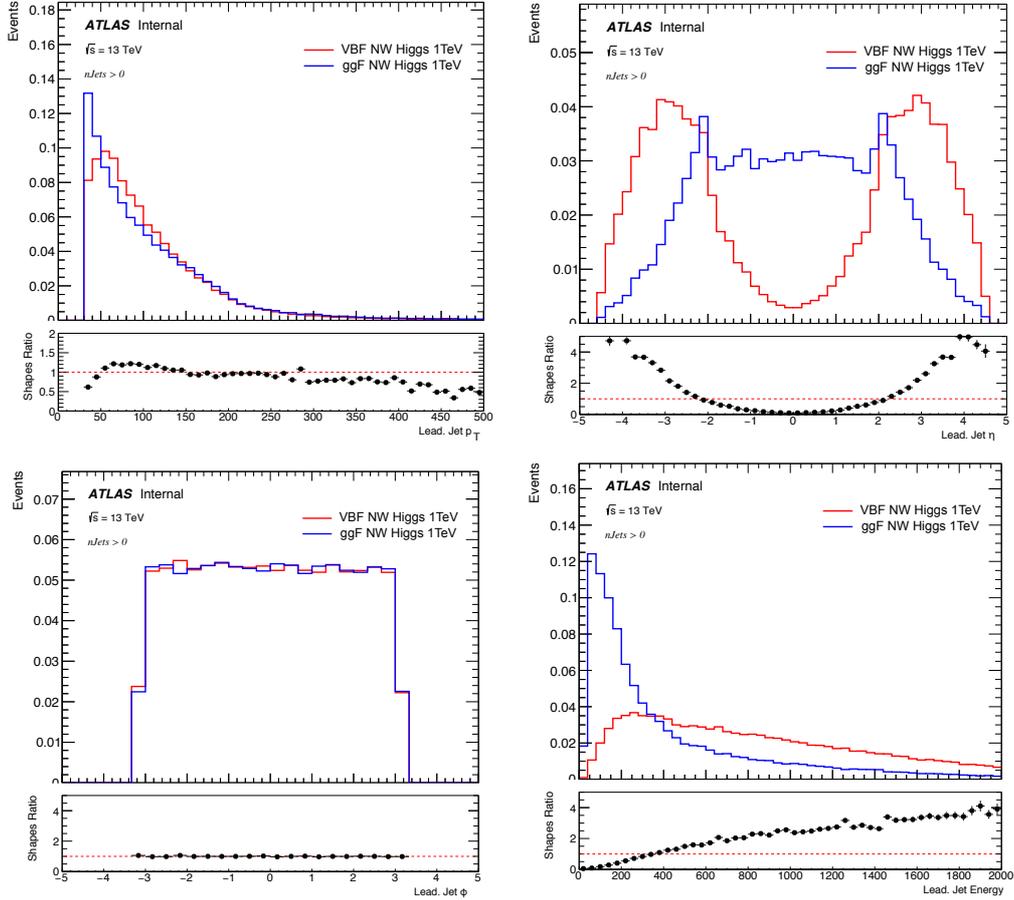


Figure 4.15: 4-momentum of the leading jet used as RNN input. The shape comparison for VBF (red) and ggF (blue) 1 TeV signals are showed.

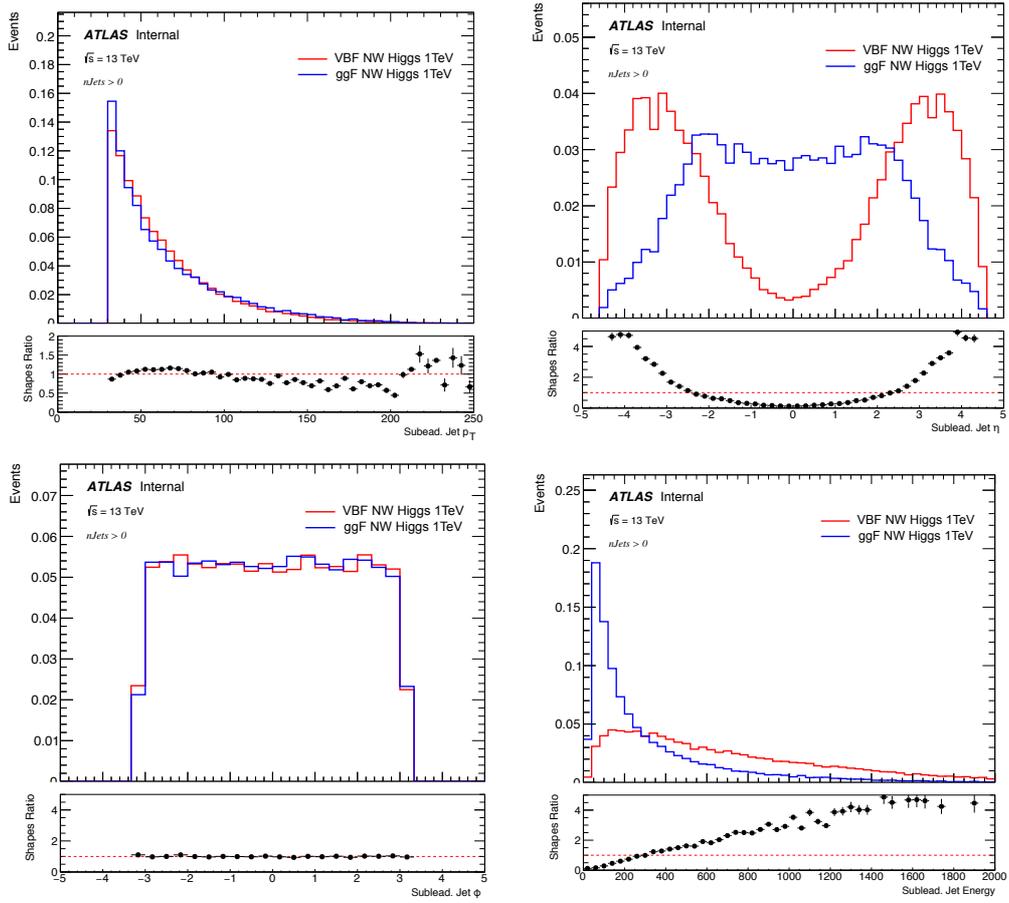


Figure 4.16: 4-momentum of the sub-leading jet used as RNN input. The shape comparison for VBF (red) and ggF (blue) 1 TeV signals are showed.

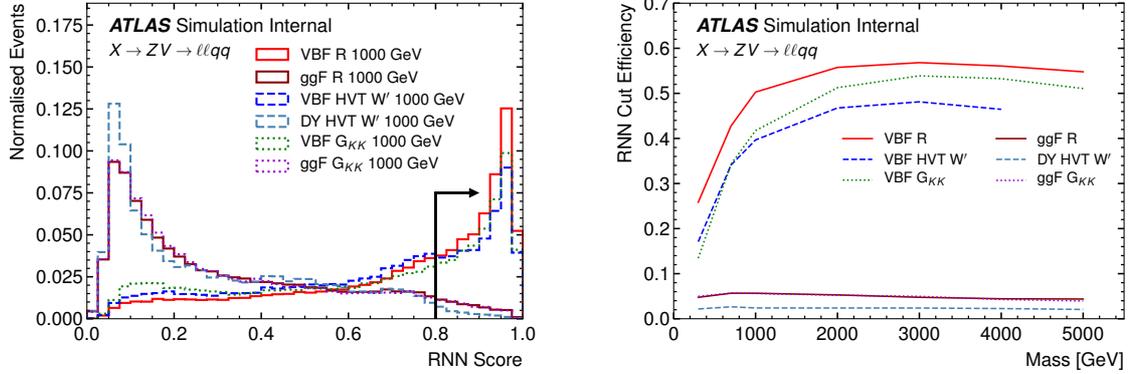


Figure 4.17: Shapes distribution of the RNN scores for different signal hypotheses (on the left) and fractions of signal events passing the VBF requirement on the RNN score as functions of the resonance mass (on the right).

#### 4.3.4 $V \rightarrow J/ V \rightarrow qq$ selection

The identification of the  $ZV \rightarrow \ell\ell qq$  decays proceeds by applying the merged  $V \rightarrow J$  selection. If this fails, the event is processed by the resolved  $V \rightarrow jj$  selection.

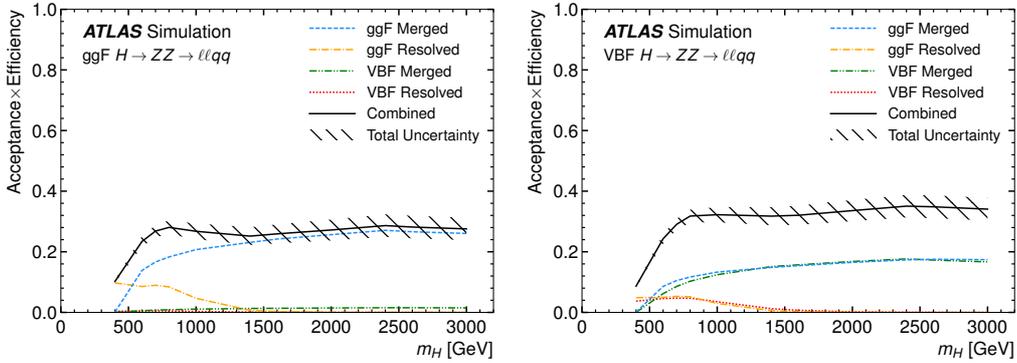


Figure 4.18: Selection acceptance times efficiency for the  $H \rightarrow ZZ \rightarrow \ell\ell qq$  events from MC simulations, as a function of the Higgs boson mass for ggF (on the left) and VBF (on the right) production, combining the merged HP and LP signal regions of the  $ZV \rightarrow \ell\ell J$  selection and the resolved regions of the  $ZV \rightarrow \ell\ell j$  selection.

The priority is optimized in order to increase the analysis sensitivity for the search of heavy resonances. Indeed a smaller background is expected in the  $ZV \rightarrow llJ$  final states. The priority is anyway important in the resonance mass region between 0.5 and 1 TeV: for resonance masses above 1 TeV the merged regime is dominant, while the resolved regime is dominant for masses below 0.5 TeV, as it is possible to see by the detector acceptance times selection efficiency plots showed in figure 4.18.

### $V \rightarrow J$ selection

For large resonance masses ( $m_X \geq 500$  GeV) the hadronic  $W/Z$  boson decays can often be reconstructed as a large-R jet. In the merged regime events from the  $Z \rightarrow ll$  candidates containing at least one large-R jet are selected, and the leading large-R jet is used to reconstruct  $W/Z$  candidates.

Selection		SR		ZCR	
		HP	LP	HP	LP
$Z \rightarrow ll$	Number of Loose leptons	2			
	Same flavor	yes			
	Leptons $p_T$	$> 30$ GeV			
	Dilepton invariant mass	$83 \text{ GeV} < m_{ee} < 99 \text{ GeV}$ $-0.01170p_T^l + 85.63 \text{ GeV} < m_{\mu\mu} < 0.01850p_T^l + 94.00 \text{ GeV}$			
	Opposite sign	For $\mu\mu$ channel only			
$W/Z \rightarrow J$	Num of large-R jets	$\geq 1$			
	$D_2$ cut	pass	fail	pass	fail
	$W/Z$ mass window cut	pass	pass	fail	fail
	Numb. of associated VR track jets $b$ -tagged	For $Z \rightarrow J$ : $\leq 1$ ( $= 2$ ) for untagged (tagged) category			
Topology cut	$\min(p_T^l, p_T^J)/m_{ZV} > 0.35(0.25)$ for the DY/ggF (VBF) category				

Table 4.7: Event selection summary for merged analysis.

In order to improve the significance of the signal over the background, events are further required to have  $\min(p_T^l, p_T^J)/m_{ZV} > 0.35$  (0.25) in the DY/ggF (VBF) category.

The  $W/Z$ -tagging, based on the large-R jet window mass and the  $D_2$  substructure variable, is required to select the  $W/Z \rightarrow q\bar{q}$  candidates: the events passing both mass and  $D_2$  cuts are selected in the High-Purity (HP) SR, while events passing the mass window cut but failing the  $D_2$  cut are selected in the Low-Purity (LP) SR.

The tagger works differently for the identification of the  $Z \rightarrow q\bar{q}$  and  $W \rightarrow q\bar{q}$  candidates, even if there is a large overlap between the selections of the large-R jet window mass and the  $D_2$ , as it is showed in figure 4.7. Depending on the spin hypothesis, the analysis looks for  $ZZ \rightarrow llqq$  (spin-0/spin-2) and  $ZW \rightarrow llqq$  (spin-1) candidates.

In the  $Z \rightarrow q\bar{q}$  selection, events are further divided into “tagged” and “untagged” categories (only for the ggF/DY category), applying  $b$ -tagging to the

two leading VR track jets ghost-associated [110] to the large-R jet:

- “Untagged”: 0 or 1  $b$ -tagged VR jets ghost-associated to the large-R jet;
- “Tagged”: 2  $b$ -tagged VR jets ghost-associated to the large-R jet.

A summary of the event selections in the merged category is reported in table 4.7.

### $V \rightarrow jj$ selection

For resonance masses  $\leq 1$  TeV, the  $W/Z \rightarrow q\bar{q}$  decays can be reconstructed as two well-separated small-R jets. In the resolved regime events from the  $Z \rightarrow ll$  candidates containing at least two small-R jet in the central region of the detector ( $|\eta| < 2.5$ ) are selected.

As in the merged regime, for the  $Z \rightarrow q\bar{q}$  selection in the ggF/DY category the candidate events are divided into “tagged” category, containing events with two  $b$ -tagged jets, and “untagged” category, containing events with less than two  $b$ -tagged jets.

In the tagged category the two  $b$ -tagged jets are chosen to reconstruct the hadronic  $Z$  decay, while in the untagged category the two jets with the highest  $p_T$  are chosen, regardless of their  $b$ -tagged status.

For the  $W \rightarrow q\bar{q}$  identification, no  $b$ -tagging selection is applied and the two signal jets with highest  $p_T$  are selected. If both selected jets are  $b$ -tagged, the event is discarded.

In both the  $Z \rightarrow q\bar{q}$  and  $W \rightarrow q\bar{q}$  selections, events with more than 2  $b$ -tagged jets are rejected in order to suppress the  $t\bar{t}$  contamination.

After selecting the two jets from the  $W/Z$  decay, the leading jet of the two is required to have  $p_T > 60$  GeV to maximize the sensitivity in the low resonance mass region, while the sub-leading jet is required to have  $p_T > 30$  GeV. Similar to the merged regime, events are required to have  $\min(p_T^l, p_T^{jj})/m_{ZV} > 0.35$  (0.25) in DY/ggF (VBF) category. Finally, to select events that are consistent with hadronically decaying  $Z$  ( $W$ ) boson, a window cut  $70(62) < m_{jj} < 105(97)$  GeV is required.

A summary of the event selections in the resolved category is reported in table 4.8.

Selection		SR	ZCR
$Z \rightarrow \ell\ell$	Number of Loose leptons	2	
	Same flavor	yes	
	Leptons $p_T$	$> 30$ GeV	
	Dilepton invariant mass	$83 \text{ GeV} < m_{ee} < 99 \text{ GeV}$ $-0.01170p_T^{\ell\ell} + 85.63 \text{ GeV} < m_{\mu\mu} < 0.01850p_T^{\ell\ell} + 94.00 \text{ GeV}$	
	Opposite sign	For $\mu\mu$ channel only	
$W/Z \rightarrow jj$	Num of signal small-R jets	2	
	Leading jet $p_T$	$> 60$ GeV	
	Subleading jet $p_T$	$> 30$ GeV	
	$Z \rightarrow q\bar{q}$	$78 \text{ GeV} < m_{jj} < 105 \text{ GeV}$	$50 \text{ GeV} < m_{jj} < 68 \text{ GeV}$ or
	$W \rightarrow q\bar{q}$	$68 \text{ GeV} < m_{jj} < 98 \text{ GeV}$	$105 \text{ GeV} < m_{jj} < 150 \text{ GeV}$
Num. of $b$ -tagged jets	For $Z \rightarrow jj$ : $\leq 1$ ( $= 2$ ) for untagged (tagged) category		
Topology cut	$\min(p_T^{\ell\ell}, p_T^{jj})/m_{ZV}$	$> 0.35(0.25)$ for the DY/ggF (VBF) category	

Table 4.8: Event selection summary for resolved analysis.

### Optimization of the $p_T$ ratio cut

Part of my work has been dedicated on the optimization of the topology cut called “ $p_T$  ratio”. In the previous analysis a different cut for merged and resolved categories, based on the event topology, was applied:

- Merged regime:  $\min(p_T^{\ell\ell}, p_T^{jj})/m_{ZV} > 0.3$  (0.35) for spin-0 (spin-1/2) interpretation;
- Resolved regime:  $\sqrt{(p_T^{\ell\ell})^2 + (p_T^{jj})^2}/m_{ZV} > 0.4$  (0.5) for spin-0 (spin-1/2) interpretation.

The choice of these ratio cuts have been revisited for the full Run-II analysis, in order to harmonize between the merged and resolved regions, signal spin hypotheses, and harmonize the selections of the  $lvqq$  and  $llqq$  lepton channels.

In the  $lvqq$  analysis, the selections are based on the minimum value of the dilepton  $p_T$  and hadronic  $p_T$  (min- $p_T$  cut variable), both in the merged and resolved regimes. So for the 2-lepton channel the gain in significance in the ggF and VBF regions has been studied, as a function of the resonance mass, for the different spin hypotheses and comparing different thresholds of the  $\min(p_T^{\ell\ell}, p_T^{jj})/m_{ZV}$  variable.

The gain is studied in the region just before the  $p_T$  ratio cut, and the following definition of significance, suitable for binned distributions, has been used [111]:

$$\sigma = \sqrt{\Sigma_i 2 \cdot [(S_i + B_i) \cdot \ln(1 + \frac{S_i}{B_i}) - S_i]} \quad (4.8)$$

where  $S_i$  and  $B_i$  are the expected number of signal and background events respectively, found in the  $i$ -th bin of the  $m_{ll}$  ( $m_{lljj}$ ) distribution in the merged (resolved) category.

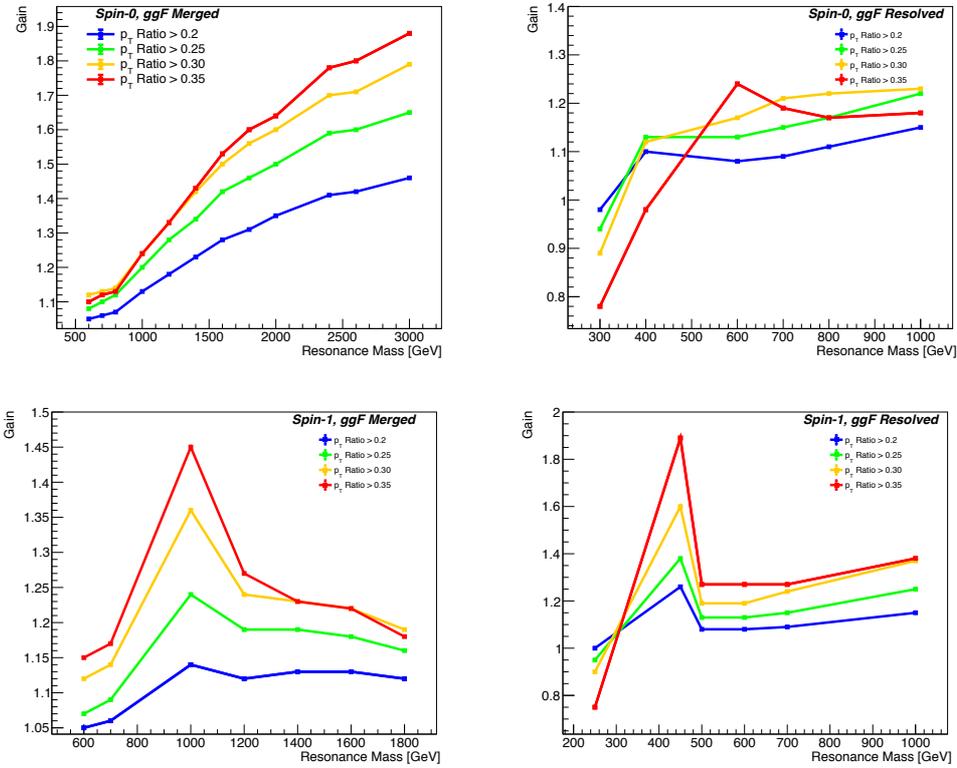


Figure 4.19: Comparisons in merged (left plots) and resolved (right plots) regime of gain in significance (significance ratio after and before cutting) for  $\min(p_T^l, p_T^{had})/m_{ZV} > 0.2$  (blue), 0.25 (green), 0.30 (orange) and 0.35 (red) for ggF spin0 NWA Higgs (upper plot) and spin 1 HVT  $W'$  (bottom plot), as a function of the resonance mass.

Figure 4.19 shows examples in the ggF merged and resolved categories for spin 0 NWA Higgs and spin 1 HVT  $W'$  signals. In the ggF regime,  $\min(p_T^l, p_T^J)/m_{ZV} > 0.35$  is the best choice that harmonize between the resolved and merged regions and for the spin 0, 1 and 2 signals. This choice

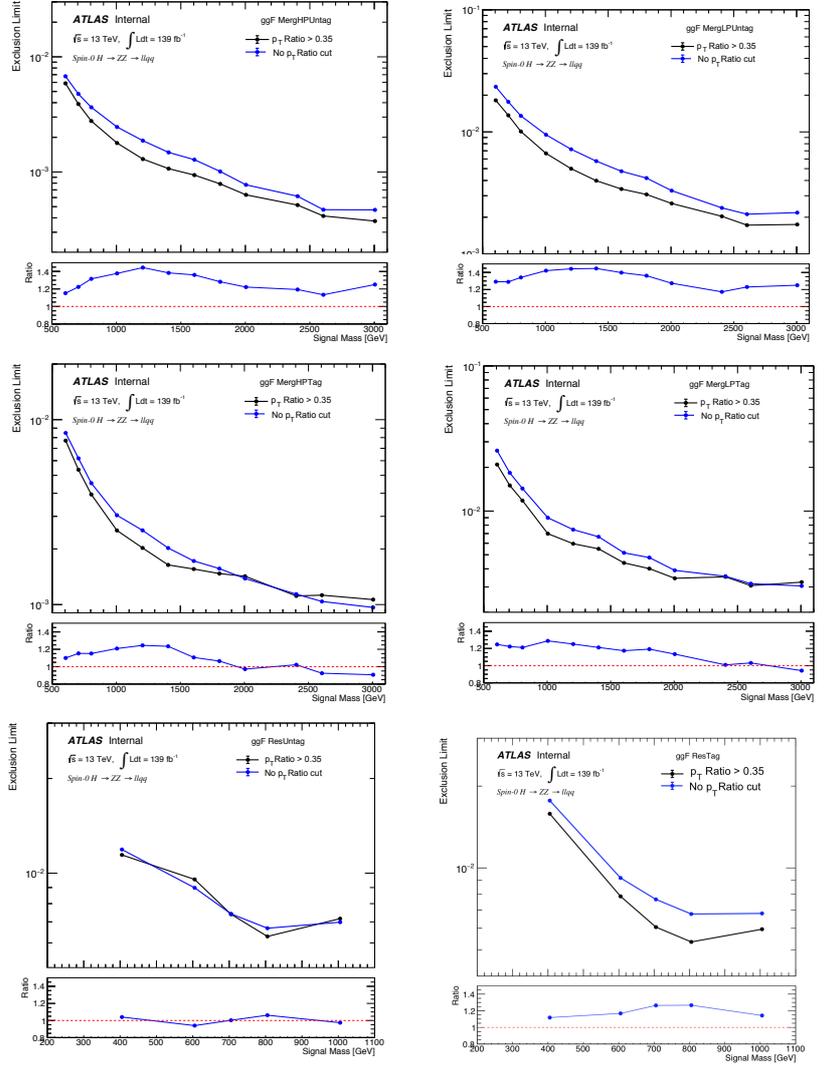


Figure 4.20: Exclusion limits with (black line) and without  $p_T$  ratio cut (blue line) greater than 0.35 for the different ggF signal regions: (a) ggF Merged HP Untagged (b) ggF Merged LP Untagged (c) ggF Merged HP Tagged (d) ggF Merged LP Tagged (e) ggF Resolved Untagged (f) ggF Resolved Tagged.

was shown to be optimal also in the analysis of the 1-lepton channel.

Figures 4.20 show exclusion limits with and without the  $p_T$  ratio cut greater than 0.35 for the different ggF 2-lepton signal regions.

The cut  $\min(p_T^I, p_T^J) / m_{ZV} > 0.35$  gives relevant improvements to the limits in all regions.

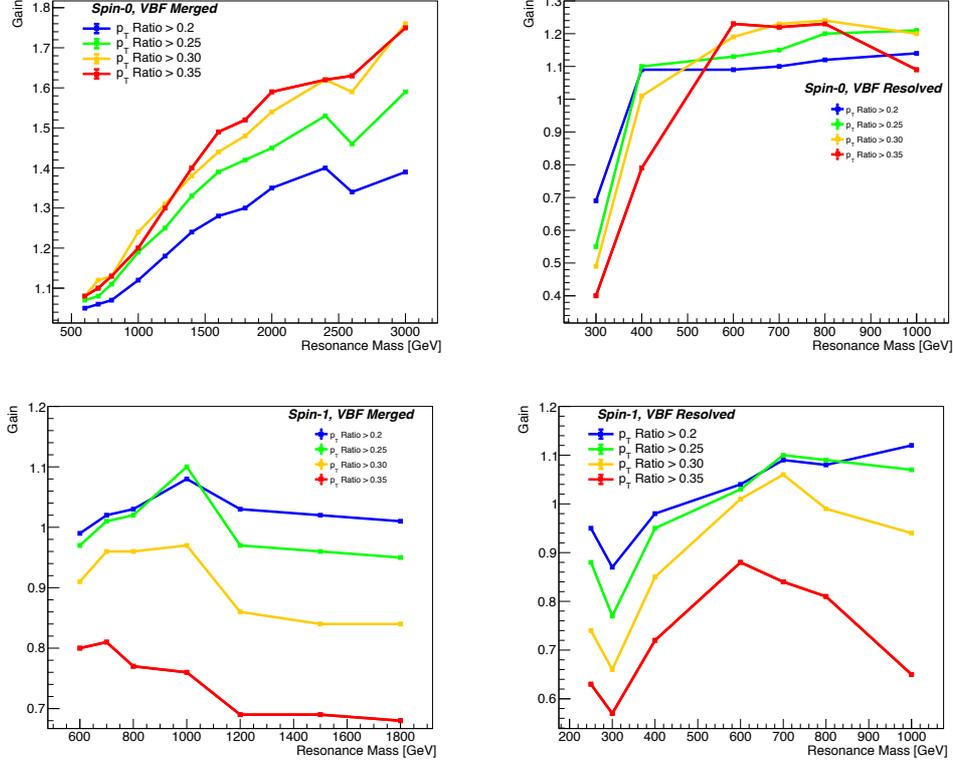


Figure 4.21: Comparisons in merged (left plots) and resolved (right plots) regime of gain in significance (significance ratio after and before cutting) for  $\min(p_T^l, p_T^{had})/m_{ZV} > 0.2$  (blue), 0.25 (green), 0.30 (orange) and 0.35 (red) for VBF spin0 NWA Higgs (upper plot) and spin 1 HVT  $W'$  (bottom plot), as a function of the resonance mass.

Figure 4.21 shows the gains in significance in the merged and resolved VBF categories. The choice  $\min(p_T^l, p_T^J)/m_{ZV} > 0.35$  is not the optimal threshold for the VBF regions.

Exclusion limits testing different  $p_T$  ratio cut hypotheses have been performed for the VBF signal regions, as showed in Figures 4.22. The choice  $\min(p_T^l, p_T^J)/m_{ZV} > 0.25$  is the best compromise between merged and resolved regime and spin 0, 1, 2 signals in the 2-lepton channel.

As a result of this study a common cut on the  $p_T$  ratio variable for spin-0, spin-1 and spin-2 signals has been applied: For Merged/Resolved ggF regions  $\min(p_T^l, p_T^J)/m_{ZV} > 0.35$  is applied. For Merged/Resolved VBF

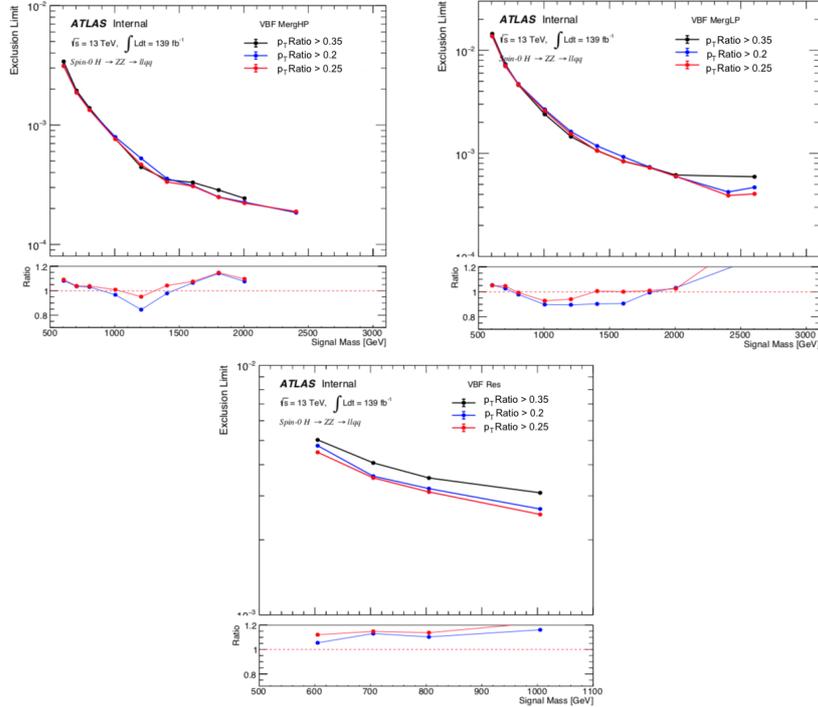


Figure 4.22: Exclusion limits with  $p_T$  ratio cut greater than 0.35 (black line), 0.25 (red line), 0.2 (blue line), for the different VBF final regions: (upper left) VBF Merged HP (upper right) VBF Merged LP (down) VBF Resolved.

regions  $\min(p_T^l, p_T^j) / m_{ZV} > 0.25$  is applied.

## 4.4 Background estimation

The main background processes for the  $X \rightarrow ZV \rightarrow llqq$  analysis are the production of leptonically decaying  $Z$  bosons in association with jets, denoted as  $Z + jets$ , Standard Model diboson production and top contributions, while leptonically decaying  $W$  bosons in association with jets have a negligible contribution in this channel.

The  $Z + jets$  corresponds to almost the totality of the background, i.e.  $\sim 90\%$ , while diboson contribution is around  $8\%$  and the top around  $2\%$ . The  $t\bar{t}$  becomes more important in the tagged regions ( $\sim 10\%$ ).

Control regions (CRs) are defined to validate the modelling of the major backgrounds and to constrain the overall contribution of the background in

the signal regions. Independent CRs for resolved and merged categories are used to account for possible mismodelling of the background depending on  $p_T$  range. Common CRs are used between the  $llqq$ ,  $lvqq$  and  $\nu\nuqq$  analyses.

$Z + jets$  CRs (ZCRs) are defined for the events that passes the nominal SRs selection but fail the di-jet mass cut. In the resolved regime di-jet mass is outside the Z mass range and it is required to be in the range  $50 \text{ GeV} < m_{jj} < 62 \text{ GeV}$  or  $105 \text{ GeV} < m_{jj} < 150 \text{ GeV}$  (see table 4.8). For the merged analysis, the leading large- $R$  jet must pass the substructure  $D_2$  cut but fail the mass window cut for the high purity category, and fail both mass and substructure cuts for the low purity category (see table 4.7). In total there are 9 ZCRs for the spin-0 interpretation ( $X \rightarrow ZZ$ ): VBF merged HP (high-purity), merged LP (low-purity) and resolved, ggF merged HP tagged/untagged, merged LP tagged/untagged and resolved tagged/untagged. For the spin-1 interpretation ( $X \rightarrow WZ$ ) the ZCRs are in total 6 since there aren't the tagged and untagged categories.

$W + jets$  and  $t\bar{t}$  are relevant background for the other leptons channels and dedicated control regions are defined in the 1-lepton analysis, while the diboson production is estimated directly from Monte Carlo.

In appendix A are summarised the informations about the cross sections of these background processes.

## ZCR plots

Comparisons of distributions of various kinematic variables, like  $p_T$ , mass,  $D_2$  of the leading large- $R$  jet and the invariant mass of the  $ZV$  system, between data and MC in different merged ZCRs are shown:

- Fig. 4.23 for the ggF merged HP untagged ZCR,
- Fig. 4.24 for the ggF merged LP untagged ZCR,
- Fig. 4.25 for the ggF merged HP tagged ZCR,
- Fig. 4.26 for the ggF merged LP tagged ZCR,
- Fig. 4.27 for the VBF merged HP ZCR,
- Fig. 4.28 for the VBF merged LP ZCR.

Similar plots are shown for the resolved ZCRs:

- Fig. 4.29 for the ggF resolved untagged ZCR,
- Fig. 4.30 for the ggF resolved tagged ZCR,
- Fig. 4.31 for the VBF resolved ZCR.

In general, the observed data agree well with MC predictions in all regions and no significant shape mismodelling is observed. A slope is generally observed on large-R jet  $p_T$  distributions, as shown in figure 4.23. Since TCC jet  $p_T$ -scale is not corrected by in-situ calibration, the slope is as expected and it is covered by the conservative uncertainty based on  $R_{trk}$  method. As shown in figures 4.25–4.26, in merged  $b$ -tagged categories, the overall normalization of the MC sample is about 20% lower than data, while it is about 5% higher than data in untagged categories in figures 4.23–4.24. The observation can be explained by the modeling of the fraction of  $Z$ +heavy-flavor jets. As reported in Ref. [112], the fraction of  $Z$ +HF in Sherpa MC can be about 20% lower than the observed data. It justifies to use of separate normalization factors for  $Z$ +jets in tagged and untagged categories in the final fit. The similar feature is observed for the resolved category. In the tagged about 20% normalization disagreement is found (figure 4.30), while good agreement is found in the untagged region (figure 4.29). In VBF categories in figures 4.27, 4.28 and 4.31, the MC simulated events overestimate the data by 5 (20%) for resolved (merged) channel. Different normalization factors for  $Z$ +jets in VBF categories are used in order to absorb possible mismodelling on the RNN score distribution.

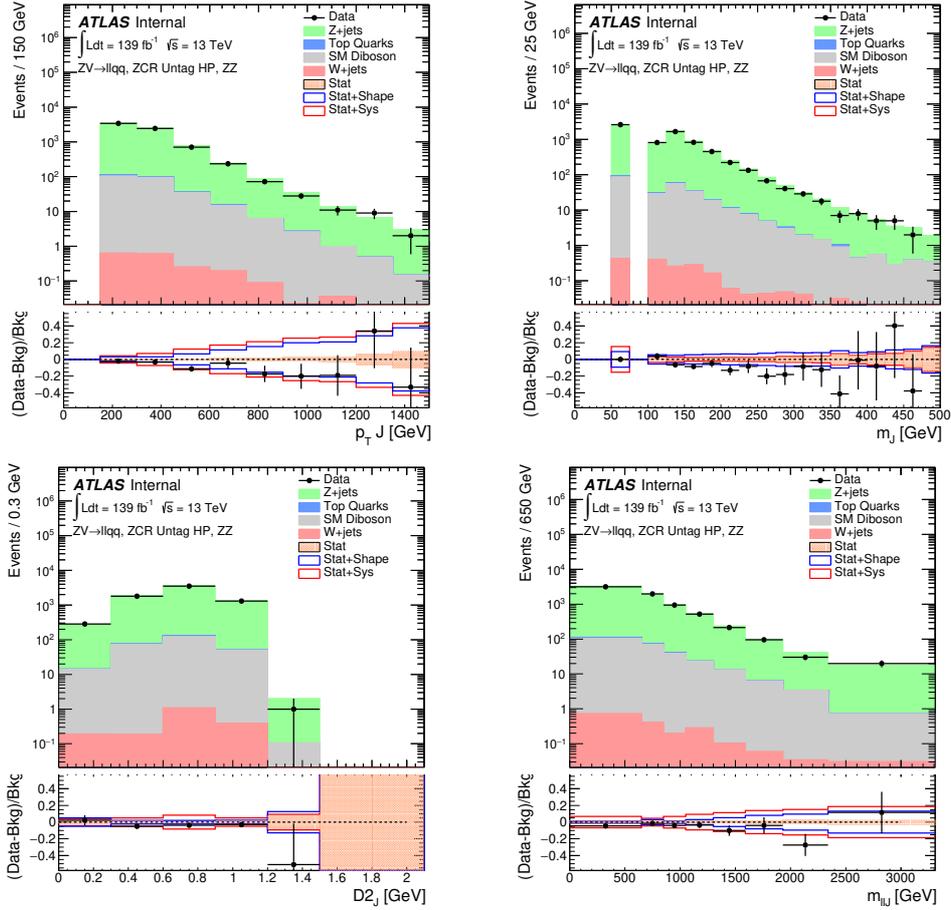


Figure 4.23: Comparisons of the observed 2015-2018 data and expected background event yields in the ggF Merged HP Untagged ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Fat jet  $p_T$  distribution. (upper right) Fat jet mass distribution. (lower left) Fat jet  $D_2$  distribution. (lower right)  $M_{llJ}$  discriminant distribution.

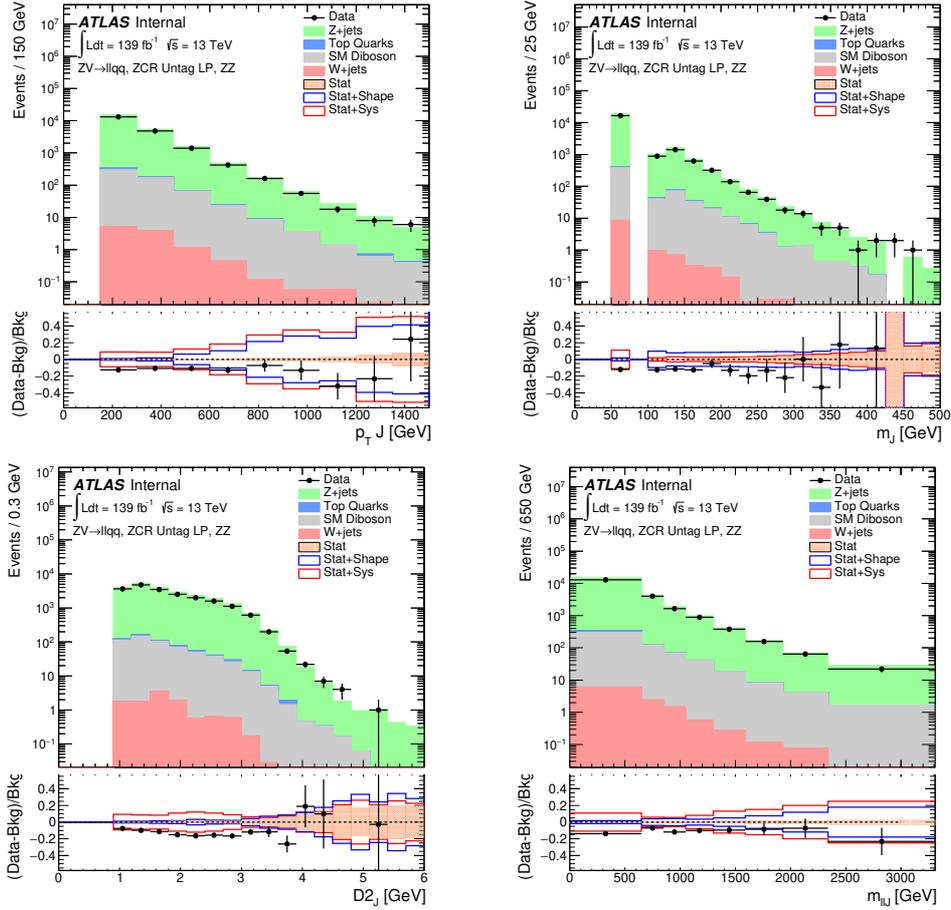


Figure 4.24: Comparisons of the observed 2015-2018 data and expected background event yields in the ggF Merged LP Untagged ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Fat jet  $p_T$  distribution. (upper right) Fat jet mass distribution. (lower left) Fat jet  $D_2$  distribution. (lower right)  $M_{llJ}$  discriminant distribution.

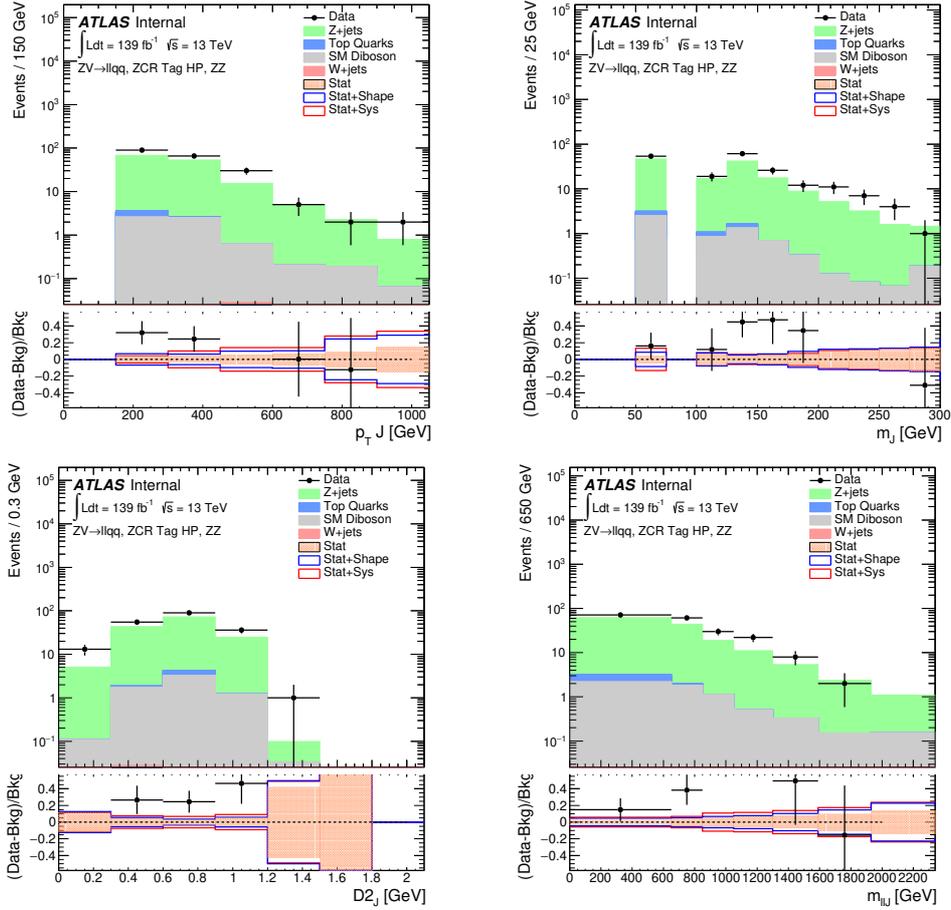


Figure 4.25: Comparisons of the observed 2015-2018 data and expected background event yields in the ggF Merged HP Tagged ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Fat jet  $p_T$  distribution. (upper right) Fat jet mass distribution. (lower left) Fat jet  $D_2$  distribution. (lower right)  $M_{llJ}$  discriminant distribution.

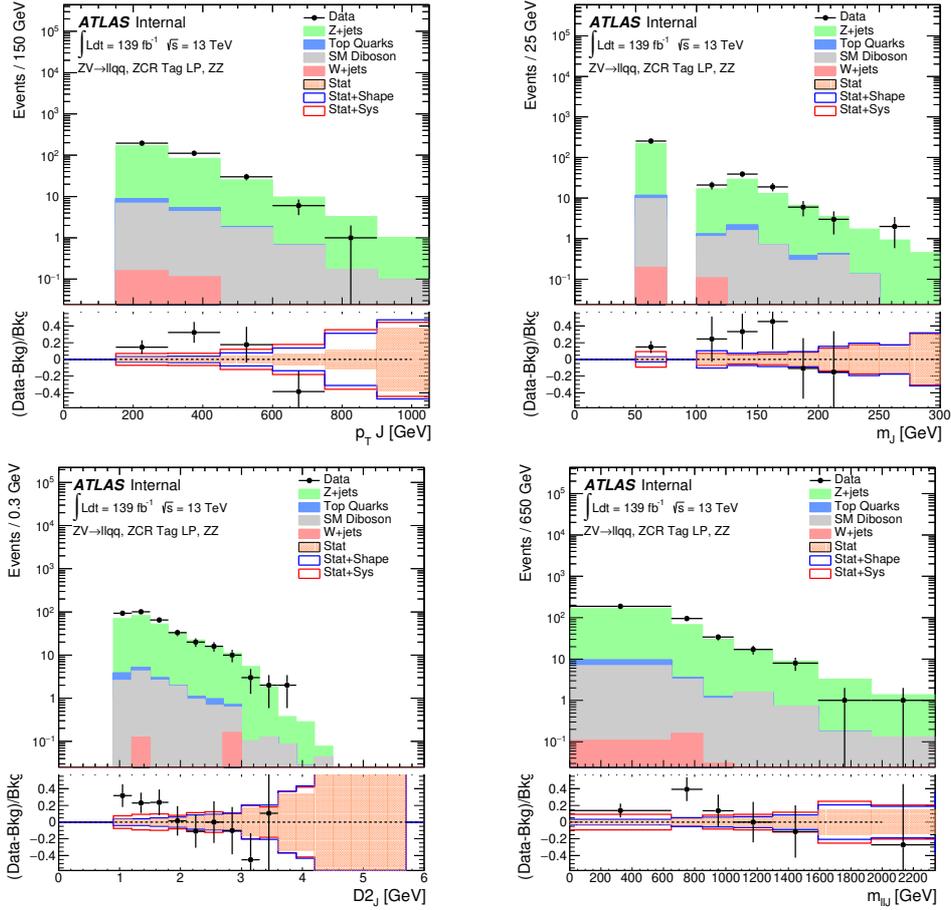


Figure 4.26: Comparisons of the observed 2015-2018 data and expected background event yields in the ggF Merged LP Tagged ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Fat jet  $p_T$  distribution. (upper right) Fat jet mass distribution. (lower left) Fat jet  $D_2$  distribution. (lower right)  $M_{llJ}$  discriminant distribution.

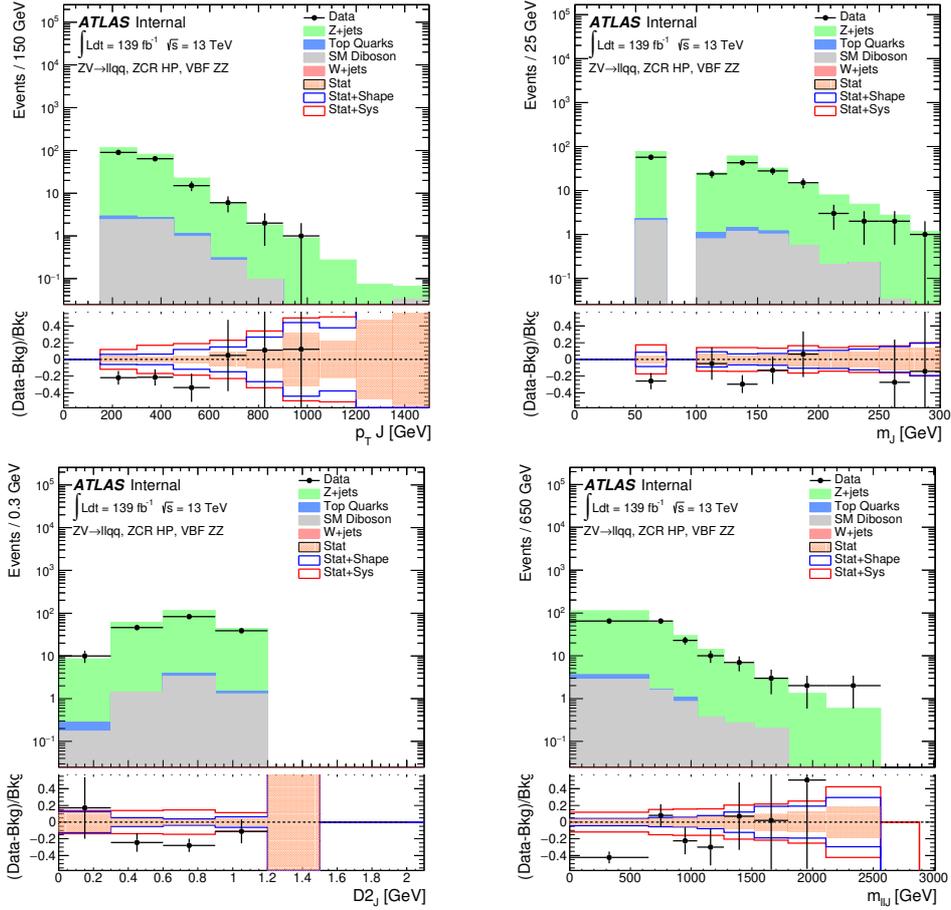


Figure 4.27: Comparisons of the observed 2015-2018 data and expected background event yields in the VBF Merged HP ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Fat jet  $p_T$  distribution. (upper right) Fat jet mass distribution. (lower left) Fat jet  $D_2$  distribution. (lower right)  $MllJ$  discriminant distribution.

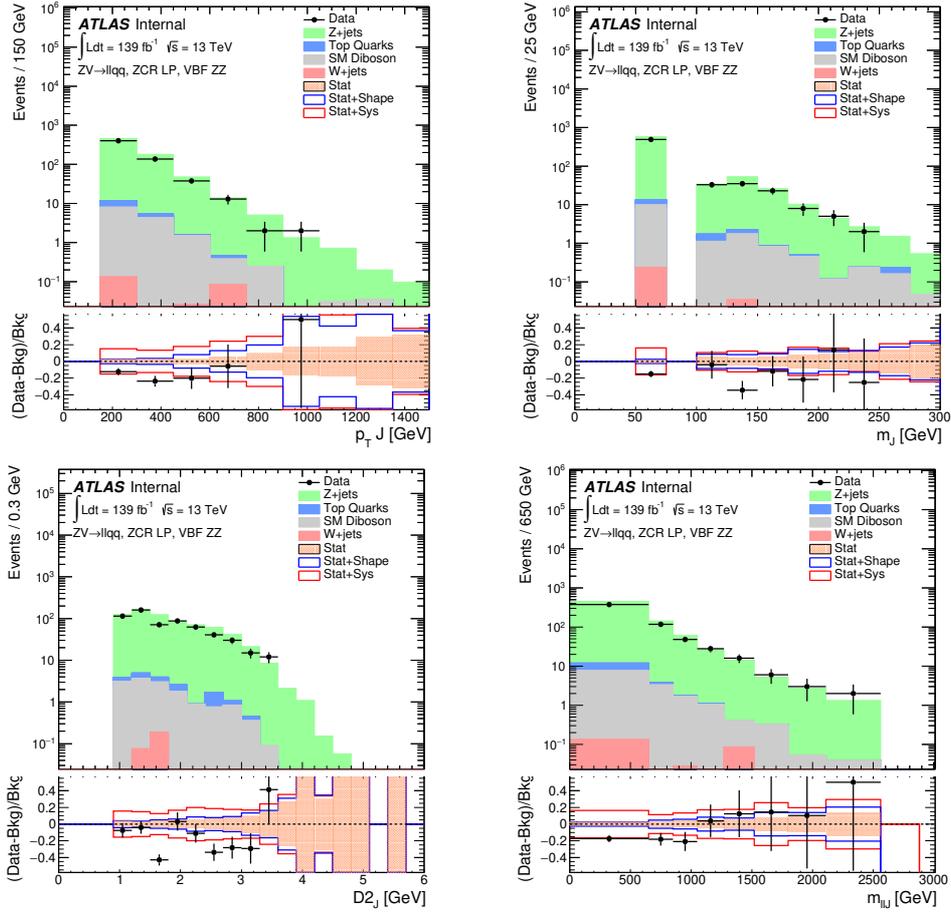


Figure 4.28: Comparisons of the observed 2015-2018 data and expected background event yields in the VBF Merged LP ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Fat jet  $p_T$  distribution. (upper right) Fat jet mass distribution. (lower left) Fat jet  $D_2$  distribution. (lower right)  $MllJ$  discriminant distribution.

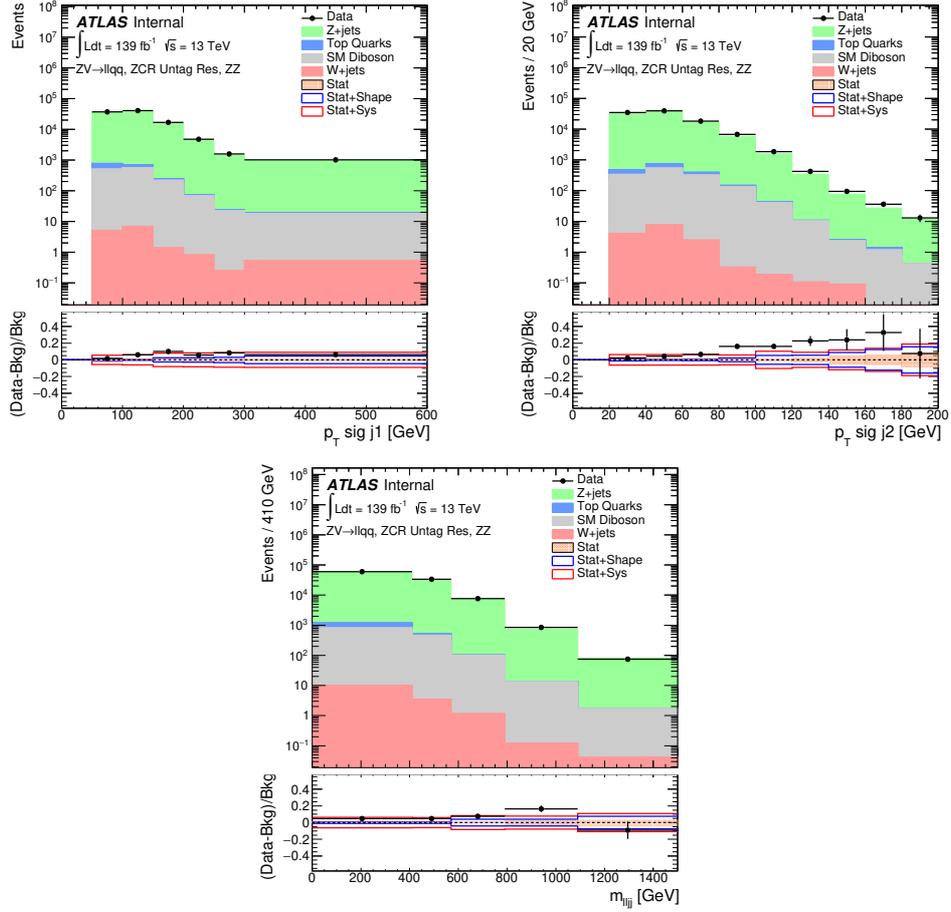


Figure 4.29: Comparisons of the observed 2015-2018 data and expected background event yields in the ggF Resolved Untagged ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Signal jet1  $p_T$  distribution. (upper right) Signal jet2  $p_T$  distribution. (lower)  $M_{lljj}$  discriminant distribution.

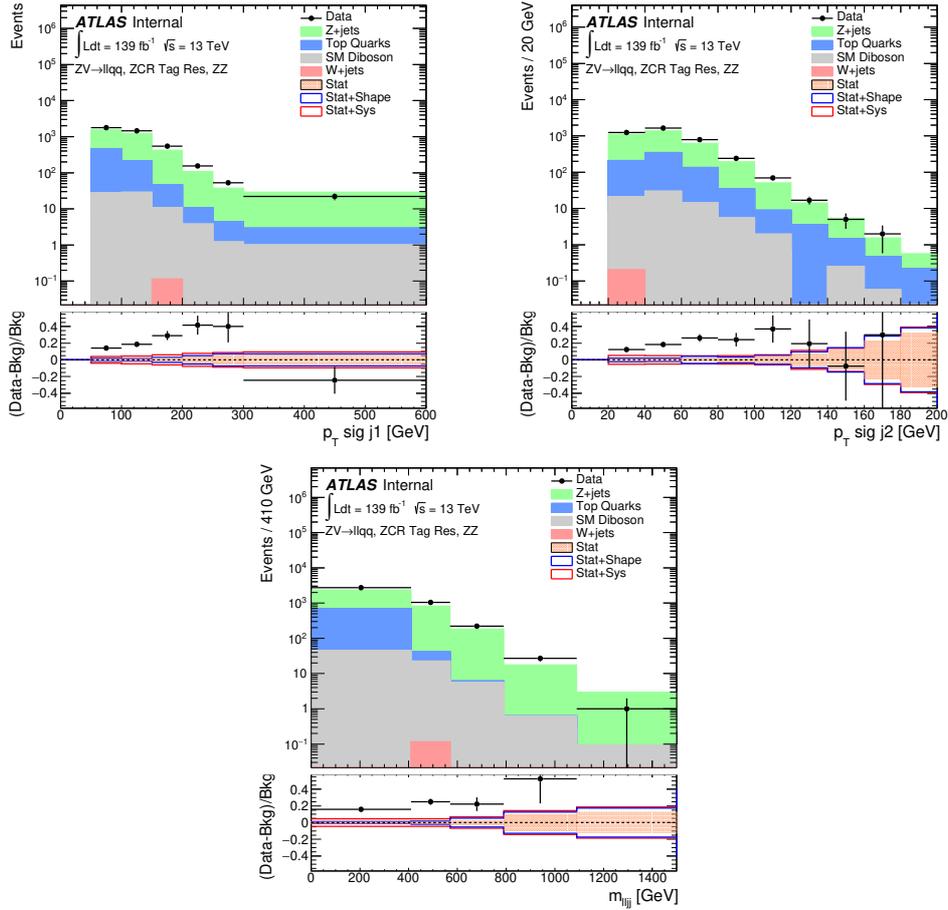


Figure 4.30: Comparisons of the observed 2015-2018 data and expected background event yields in the ggF Resolved Tagged ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Signal jet1  $p_T$  distribution. (upper right) Signal jet2  $p_T$  distribution. (lower)  $M_{lljj}$  discriminant distribution.

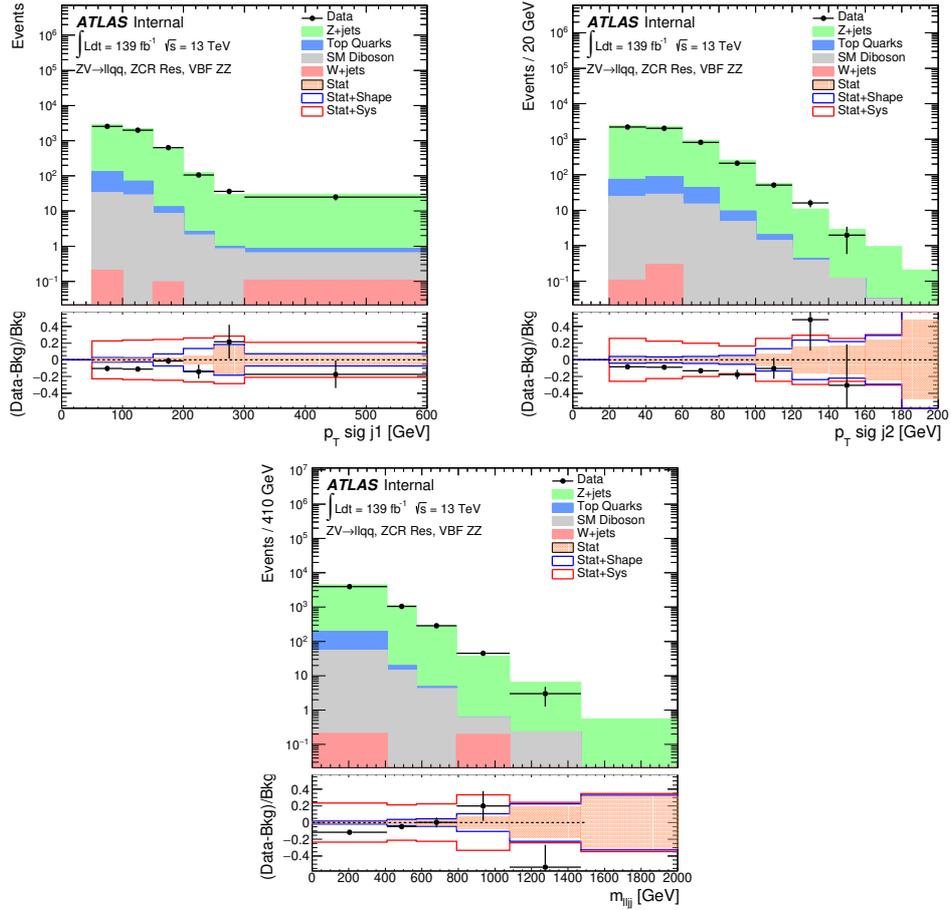


Figure 4.31: Comparisons of the observed 2015-2018 data and expected background event yields in the VBF Resolved ZCR of the  $H \rightarrow ZZ \rightarrow llqq$  search: (upper left) Signal jet1  $p_T$  distribution. (upper right) Signal jet2  $p_T$  distribution. (lower)  $M_{lljj}$  discriminant distribution.

## 4.5 Systematic uncertainties

Systematics uncertainties give a measure of the impact of experimental and theoretical uncertainties to the final result. They are divided in three categories: experimental uncertainties, uncertainties on the modelling of background processes and theoretical uncertainties on the signal processes. Each systematic uncertainty is treated as a nuisance parameter and their variations are estimated on the final discriminant.

### 4.5.1 Experimental uncertainties

#### Luminosity

The uncertainty on the integrated luminosity for the 2015+2016 dataset is 2.1%, while is 2.4% for the 2017 and 2% for the 2018 dataset. The uncertainty for the combined Run-2 dataset is 1.7%. The luminosity uncertainty is applied to backgrounds and signal Monte Carlo samples.

#### Pileup reweighting

An uncertainty associated with the pileup reweighting is also considered: a variation in the pile-up reweighting of MC events is included to cover the uncertainty on the ratio between the predicted and measured inelastic cross section, as reported in Ref. [113].

#### Trigger

Systematic uncertainties on the efficiency of the electron or muon triggers are evaluated using the tag and probe method. It is applied to backgrounds and signal Monte Carlo samples.

#### Muons and electrons

Both electrons and muons have different sources of uncertainties, and each of them is treated separately. The following systematic uncertainties are applied:

- Identification and reconstruction efficiencies;
- Isolation efficiency;
- Energy and momentum scales.

These efficiencies and the related uncertainties are measured by tag and probe method using the  $Z$  mass peak. Moreover for the muons the additional track-to-vertex association efficiency is considered.

### Small- $R$ and Large- $R$ Jet Energy Scale and resolution uncertainty

The jet energy scale and resolution of the small- $R$  jets are measured in-situ by calculating the response between MC and data in various bins of kinematic phase space.

The  $p_T$  of the TCC jet can be approximated to be equivalent with calo-jet  $p_T$ , and not depending significantly on track measurements. The uncertainty on the  $p_T$  scale of TCC jet is evaluated with the  $R_{trk}$  method, by comparing the ratio of TCC jet  $p_T$  to track-jet  $p_T$ , in dijet data and simulation. In addition to this uncertainty are also considered the uncertainties on track measurements, the difference between Pythia and Sherpa di-jet simulations and the statistical uncertainty of di-jet data. In order to take into account the possible difference between the  $p_T$  of TCC and calo-jet, the difference between TCC jet and calo-jet  $p_T$  in di-jet data and simulation is additionally considered.

The large- $R$  jet resolution uncertainty is evaluated event-by-event by rerunning the analysis with an additional Gaussian smearing, with 2% width, applied to the large- $R$  jet  $p_T$  to degrade the nominal resolution by the systematic uncertainty value.

### $W/Z$ -tagging efficiency scale factor (SF) uncertainty

The  $p_T$  of the TCC jet is equivalent to the one of the LCTopo jet in the first approximation, so the  $R_{trk}$  method can be used in the estimation of the scale uncertainty. However large- $R$  jet  $D_2$  and mass highly depend on the tracks informations due to the TCCs algorithm. So the efficiency of  $W/Z$ -tagging, based on the large- $R$  jet mass and  $D_2$  variables, and its uncertainty are estimated in data using  $t\bar{t}$  control sample and corrected by taking the ratio of efficiencies between data and Monte Carlo. The efficiency to single- $q/g$  background is instead estimated using di-jet sample.

The efficiency scale factor and uncertainties are estimated in each of the 4 different regions:

- Pass mass window and  $D_2$  cut (HP SR);
- Pass mass window and fail  $D_2$  cut (LP SR);
- Fail mass window and pass  $D_2$  cut (HP CR);

- Fail mass window and  $D_2$  cut (LP CR);

which are used to define SR and CR in the analysis, and correlations between 4 regions are correctly taken into account.

### **B-tagging systematics**

The systematic uncertainties associated to the  $b$ -tagging are evaluated as uncertainties on the scaling factor to take account for possible disagreement of the  $b$ -tag efficiency between data and MC. Separated scale factors and corresponding systematic uncertainties are provided for  $b$ -,  $c$ - and light-flavor-induced jets, based on several measurements.

## **4.5.2 Background modeling uncertainties**

Different kind of systematics have been evaluated to take into account the uncertainties in the modelling of backgrounds. These include shape systematic on the final discriminant, as well as a normalization systematic for predictions estimated from simulation only.

### **$W/Z$ +jets modeling uncertainty**

Several sources of modelling uncertainties may affect the shape of the  $W/Z$ +jets samples in both the resolved and merged channels. Scale variations of the re-normalization, factorization scale and variations of the PDF are taken into account. Additionally, a matrix element and parton shower variations have been considered by comparing the nominal Sherpa samples with MadGraph+Pythia samples.

The overall normalization of the  $W/Z$ +jet is a free parameter in the fit and thus only shape variations of  $W/Z$ +jet distributions are considered.

### **$t\bar{t}$ modeling uncertainty**

An uncertainty on the shape of the final discriminant used in the fit for  $t\bar{t}$  is derived comparing the default POWHEG sample with the distribution obtained using AMC@NLO as an alternative generator. The changes between different generators are evaluated by taking the ratio of the shape systematic to the base MC sample.

Additional systematic uncertainties are evaluated by comparing the nominal sample showered with Pythia to one showered with Herwig. Sample with the scale doubled and halved are also used for the evaluation of the systematic

uncertainty, by taking the ratio between each variation sample and the corresponding nominal sample.

The overall normalization of the  $t\bar{t}$  is a free parameter in the fit.

### Diboson and single top modeling uncertainty

The Monte Carlo simulation is used to estimate the diboson background, since it is not possible to define a dedicated control region to constrain its modelling. The normalization uncertainty in diboson prediction is estimated using Sherpa samples, which were generated with up to one additional parton at NLO and up to three partons at LO.

The uncertainties considered are scale variations of the re-normalization, factorization scale and variations of the PDF. A shape systematic is extracted using alternative generators (Powheg+Pythia) instead of the nominal Sherpa. Single-t process is a minor component in the analysis, and the pure MC prediction for the shape with conservative uncertainty of 20% have been used.

### 4.5.3 Signal theoretical uncertainties

Additional systematics are introduced due to modelling differences between various signal Monte Carlo generators:

- QCD scale uncertainties: They are estimated for signal samples generated at NLO QCD. The effect of the QCD scale uncertainty on the signal acceptance is estimated by varying the factorization and re-normalization scales.
- PDF uncertainties: They are estimated by taking the acceptance difference due to internal PDF error sets and the difference between choice of PDF sets. For HVT, Graviton and Radion signal samples, NNPDF23LO PDF set is used, which has an ensemble of 100 error PDFs. Alternative PDF sets of Cteq6L1 and MSTW2008LO are used to assess the PDF uncertainties.
- Parton shower uncertainty: In order to evaluate the Initial State Radiation (ISR)/ Final State Radiation (FSR) uncertainty, a variation of the relevant parameters in Pythia8 is performed.

## 4.6 Statistical approach and results

The statistical treatment consists in a combined profile likelihood fit to binned discriminants in all categories and regions simultaneously. It is based

on the framework described in [114, 115, 116].

The three research channels  $X \rightarrow ZV \rightarrow llqq$ ,  $X \rightarrow ZV \rightarrow \nu\nu qq$  and  $X \rightarrow WV \rightarrow l\nu qq$  have been combined together to gain in sensitivity. A binned profile likelihood fit to the final discriminants, that are the transverse mass  $m_T$  in the 0-lepton channel,  $m_{l\nu J}/m_{l\nu jj}$  in the 1-lepton channel and  $m_{llJ}/m_{lljj}$  in the 2-lepton channel is performed.

The binning of the discriminants for each region is determined such that the bins are larger than the signal mass resolution, with a requirement that the number of background events in the bin is non-zero, and that the relative statistical uncertainty on the number of background events in the bin is less than 70%.

The processes considered in the fit are:

- Resonance signals;
- $W + jets$ ;
- $Z + jets$ ;
- $t\bar{t}$ ;
- single-top:  $s$ -,  $t$ -,  $Wt$ -channel;
- Diboson contribution: QCD production of  $WW/WZ/ZZ$  processes;
- Multijet contribution: estimated with a data-driven method, used only in the resolved regime of the 1-lepton channel.

Four different fit configurations are considered, to search for different signal models:

- $ZZ + WW$ , to search for heavy neutral scalars and spin-2 Gravitons;
- $ZW$ , to probe for heavy charged spin-1 vectors ( $W'^{\pm}$ );
- $WW$ , to probe for heavy neutral spin-1 vectors ( $Z'$ );
- $ZW + WW$ , to probe for heavy vector triplets ( $W'^{\pm}, Z'$ ).

Several SRs and CRs of the three lepton channels are fitted simultaneously. The regions used in each fit model are summarized in tables 4.9, 4.10 4.11 and 4.12.

The binned likelihood function  $\mathcal{L}(\mu, \vec{\theta})$  depends on the signal-strength parameter  $\mu$ , a multiplicative factor applied to the theoretical signal production

Regions		Spin-0 and Spin-2 fit model ( $X \rightarrow ZZ/WW$ )						
		Merged HP		Merged LP		Resolved		
		Tagged	Untagged	Tagged	Untagged	Tagged	Untagged	
0-lepton	ggF	SR	Shape	Shape	Shape	Shape	–	
	VBF	SR	Shape		Shape		–	
1-lepton	ggF	SR	–	Shape	–	Shape	–	Shape
		WCR	–	One bin	–	One bin	–	One bin
		TopCR	–	One bin	–	One bin	–	One bin
	VBF	SR	Shape		Shape		Shape	
		WCR	One bin		One bin		One bin	
		TopCR	One bin		One bin		One bin	
2-lepton	ggF	SR	Shape	Shape	Shape	Shape	Shape	
		ZCR	One bin	One bin	One bin	One bin	One bin	
	VBF	SR	Shape		Shape		Shape	
		ZCR	One bin		One bin		One bin	

Table 4.9: Summary of the regions entering the likelihood of the fit model to search for heavy resonances decaying to  $ZZ$  and  $WW$  in semileptonic final states. “Shape” indicates that the shape of the final discriminant is taken into account in the fit. “One bin” implies that a single bin without any shape information is used in the corresponding fit region and “–” means that the region is not present in the fit.

cross section, and on a set of nuisance parameters  $\vec{\theta}$  that describe the effects of systematic uncertainties in the signal and expected backgrounds. The signal strength of the hypothetical signal processes is assigned as the  $POI$  (parameter of interest), while the systematic uncertainties, described in the previous section, and the background normalization factors are assigned as  $NPs$  (nuisance parameters). The  $NPs$  are either free to float, or constrained using Gaussian terms defined by external studies.

The  $Z + jets$ ,  $W + jets$  and  $t\bar{t}$  control regions are included in the fit with one bin per region, that means using only their event-count information. Background contributions, including their shapes in the signal regions, are taken from MC simulations but they are allowed to vary independently within their uncertainties in each bin. Moreover, normalisation scale factors ( $SFs$ ) are applied to the MC estimates of the  $Z + jets$ ,  $W + jets$  and  $t\bar{t}$  contributions. The  $SFs$  are parameters free to float in the fit and are therefore constrained by the data in both the signal and control regions.

In tables 4.13 and 4.14 are reported the overall  $Z + jets$ ,  $W + jets$  and  $t\bar{t}$  production normalization factors used in the ggF  $WZ$  and  $WW/ZZ$  regions, and in the VBF  $WZ$  and  $WW/ZZ$  regions. The diboson contribution is constrained to the theoretical estimate within the corresponding uncertainties.

Regions		Spin-1 fit model ( $X \rightarrow ZW$ )					
		Merged HP		Merged LP		Resolved	
		Tagged	Untagged	Tagged	Untagged	Tagged	Untagged
0-lepton	ggF	SR	–	Shape	–	Shape	–
	VBF	SR	Shape		Shape		–
1-lepton	ggF	SR	Shape	Shape	Shape	Shape	Shape
		WCR	One bin	One bin	One bin	One bin	One bin
		TopCR	One bin	One bin	One bin	One bin	One bin
	VBF	SR	Shape		Shape		Shape
		WCR	One bin		One bin		One bin
		TopCR	One bin		One bin		One bin
2-lepton	ggF	SR	–	Shape	–	Shape	–
		ZCR	–	One bin	–	One bin	–
	VBF	SR	Shape		Shape		Shape
		ZCR	One bin		One bin		One bin

Table 4.10: Summary of the regions entering the likelihood of the fit model to search for heavy resonances decaying to  $ZW$  in semileptonic final states. “Shape” indicates that the shape of the final discriminant is taken into account in the fit. “One bin” implies that a single bin without any shape information is used in the corresponding fit region and “–” means that the region is not present in the fit.

Regions		Spin-1 fit model ( $X \rightarrow WW$ )					
		Merged HP		Merged LP		Resolved	
		Tagged	Untagged	Tagged	Untagged	Tagged	Untagged
1-lepton	ggF	SR	–	Shape	–	Shape	–
		WCR	–	One bin	–	One bin	–
		TopCR	–	One bin	–	One bin	–
	VBF	SR	Shape		Shape		Shape
		WCR	One bin		One bin		One bin
		TopCR	One bin		One bin		One bin

Table 4.11: Summary of the regions entering the likelihood of the fit model to search for heavy resonances decaying to  $WW$  in semileptonic final states. “Shape” indicates that the shape of the final discriminant is taken into account in the fit. “One bin” implies that a single bin without any shape information is used in the corresponding fit region and “–” means that the region is not present in the fit.

#### 4.6.1 Likelihood function and test statistic

The final discriminants distributions of the analyses are arranged as the product of channels, regimes and regions. Channel refers to the analysis with dif-

Regions		Spin-1 triplet fit model ( $W' \rightarrow ZW$ and $Z' \rightarrow WW$ )					
		Merged HP		Merged LP		Resolved	
		Tagged	Untagged	Tagged	Untagged	Tagged	Untagged
0-lepton	ggF	SR	Shape	Shape	Shape	Shape	–
	VBF	SR	Shape		Shape		–
1-lepton	ggF	SR	Shape	Shape	Shape	Shape	Shape
		WCR	One bin	One bin	One bin	One bin	One bin
		TopCR	One bin	One bin	One bin	One bin	One bin
	VBF	SR	Shape		Shape		Shape
		WCR	One bin		One bin		One bin
		TopCR	One bin		One bin		One bin
2-lepton	ggF	SR	Shape	Shape	Shape	Shape	Shape
		ZCR	One bin	One bin	One bin	One bin	One bin
		TopCR	–	–	–	–	One bin
	VBF	SR	Shape		Shape		Shape
		ZCR	One bin		One bin		One bin

Table 4.12: Summary of the regions entering the likelihood of the fit model to search for heavy triplets decaying to  $ZW$  and  $WW$  in semileptonic final states. “Shape” indicates that the shape of the final discriminant is taken into account in the fit. “One bin” implies that a single bin without any shape information is used in the corresponding fit region and “–” means that the region is not present in the fit.

ferent number of leptons in the final states (0-lepton, 1-lepton and 2-lepton). Regime refers to the merged and resolved jet reconstruction and region refers to signal and control regions. Moreover, the fit is performed separately between the VBF or ggF signal production mechanism categories. The binned likelihood function is constructed as the product of Poisson probability terms:

$$\mathcal{L}(\mu, \vec{\theta}) = \prod_c \prod_i \text{Pois} \left( n_{c_i}^{\text{obs}} | n_{c_i}^{\text{sig}}(\mu, \vec{\theta}) + n_{c_i}^{\text{bkg}}(\vec{\theta}) \right) \prod_k f_k(\theta'_k | \theta_k), \quad (4.9)$$

where the signal strength parameter  $\mu$  multiplies the expected signal yield for a given benchmark set of parameters  $n_{c_i}^{\text{sig}}$  in each histogram bin  $i$  of the discriminant from channel  $c$ , and  $n_b^{\text{bkg}}$  represents the background content for bin  $i$  from channel  $c$ . The dependence of the signal and background predictions on the systematic uncertainties is described by the set of NPs  $\vec{\theta}$ , which are parametrized by Gaussian or log-normal (for normalization uncertainties in order to maintain a positive likelihood) priors.

The expected numbers of signal and background events in each bin are func-

Normalization	ggF WZ	ggF WW/ZZ
Z+jets ggF Merg HP Tag	-	$1.30 \pm 0.14$
Z+jets ggF Merg HP Untag	$0.962 \pm 0.012$	$0.963 \pm 0.012$
Z+jets ggF Merg LP Tag	-	$1.20 \pm 0.09$
Z+jets ggF Merg LP Untag	$0.893 \pm 0.007$	$0.889 \pm 0.007$
Z+jets ggF Res Tag	-	$1.26 \pm 0.03$
Z+jets ggF Res Untag	$1.060 \pm 0.004$	$1.060 \pm 0.004$
W+jets ggF Merg HP Tag	$0.92 \pm 0.11$	-
W+jets ggF Merg HP Untag	$0.950 \pm 0.008$	$0.934 \pm 0.008$
W+jets ggF Merg LP Tag	$0.92 \pm 0.07$	-
W+jets ggF Merg LP Untag	$0.879 \pm 0.005$	$0.886 \pm 0.005$
W+jets ggF Res Tag	$1.21 \pm 0.09$	-
W+jets ggF Res Untag	$1.010 \pm 0.006$	$1.020 \pm 0.006$
$t\bar{t}$ ggF Merg HP Tag	$0.96 \pm 0.03$	-
$t\bar{t}$ ggF Merg HP Untag	$0.896 \pm 0.015$	$0.942 \pm 0.019$
$t\bar{t}$ ggF Merg LP Tag	$0.98 \pm 0.03$	-
$t\bar{t}$ ggF Merg LP Untag	$0.896 \pm 0.013$	$0.9080 \pm 0.0016$
$t\bar{t}$ ggF Res Tag	$1.000 \pm 0.010$	-
$t\bar{t}$ ggF Res Untag	$0.968 \pm 0.017$	$0.949 \pm 0.015$

Table 4.13: Fitted values of background normalization factors in the ggF WZ and WW/ZZ.

Normalization	ggF WZ	ggF WW/ZZ
Z+jets VBF Merg HP Untag	$0.94 \pm 0.06$	$0.91 \pm 0.06$
Z+jets VBF Merg LP Untag	$0.88 \pm 0.04$	$0.85 \pm 0.04$
Z+jets VBF Res Untag	$0.936 \pm 0.017$	$0.928 \pm 0.017$
W+jets VBF Merg HP Untag	$0.91 \pm 0.06$	$0.93 \pm 0.06$
W+jets VBF Merg LP Untag	$0.92 \pm 0.04$	$0.91 \pm 0.04$
W+jets VBF Res Untag	$0.95 \pm 0.02$	$0.93 \pm 0.02$
$t\bar{t}$ VBF Merg HP Untag	$0.93 \pm 0.07$	$0.85 \pm 0.06$
$t\bar{t}$ VBF Merg LP Untag	$0.69 \pm 0.06$	$0.77 \pm 0.07$
$t\bar{t}$ VBF Res Untag	$0.99 \pm 0.04$	$0.97 \pm 0.04$

Table 4.14: Fitted values of background normalization factors in the VBF WZ and WW/ZZ.

tions of  $\vec{\theta}$ . The term  $f_k(\theta'_k|\theta_k)$  represents the set of constraints on  $\vec{\theta}$  from auxiliary measurements  $\theta'_k$ : these constraints include normalisation and shape uncertainties in the signal and background models, and include the statistical uncertainties of the simulated bin content.

The nominal fit result in terms of  $\mu$  and  $\sigma_\mu$  is obtained by maximising the

likelihood function with respect to all parameters.

The test statistic  $q_\mu$  is constructed according to the profile likelihood:  $q_\mu = 2 \ln(\mathcal{L}(\mu, \hat{\theta}_\mu) / \mathcal{L}(\hat{\mu}, \hat{\theta}))$ , where  $\hat{\mu}$  and  $\hat{\theta}$  are the parameters that maximise the likelihood (with the constraint  $0 \leq \hat{\mu} \leq \mu$ ), and  $\hat{\theta}_\mu$  are the nuisance parameter values that maximise the likelihood for a given  $\mu$ . This test statistic is used to measure the compatibility of the background-only model with the observed data and for exclusion intervals derived with the  $CL_s$  method [117, 118]. The limit set on  $\mu$  is then translated into a limit on the signal cross section times branching ratio, using the theoretical cross section and branching ratio for the given signal model.

## 4.6.2 Results

The observed distributions of the final discriminants are compared with the background estimates for all the signal regions defined in the three lepton channels. The  $m_{lJ}$  and  $m_{ljj}$  distributions for all the merged and resolved SRs defined in the  $X \rightarrow ZV \rightarrow llqq$  search are reported in:

1.  $X \rightarrow ZZ$  interpretation:
  - Figure 4.32: ggF Merged SRs;
  - Figure 4.33: ggF Resolved SRs;
  - Figure 4.34: VBF Merged and Resolved SRs.
2.  $X \rightarrow WZ$  interpretation:
  - Figure 4.35: ggF Merged and Resolved SRs;
  - Figure 4.36: VBF Merged and Resolved SRs.

The data are reasonably well reproduced by the background contributions in all the final discriminant distributions. The backgrounds are estimated from a simultaneous background-only fit to the SRs and their corresponding CRs.

No significant excess has been observed in any of the three lepton channels, as reported in tables 4.15-4.16 where the post-fit estimated background events from different sources in all  $WW/ZZ$  and  $WZ$  SRs are compared with the numbers of the observed events in data.

In the absence of significant deviations from expected background, constraints on the production cross sections of heavy resonances decaying to  $VV$

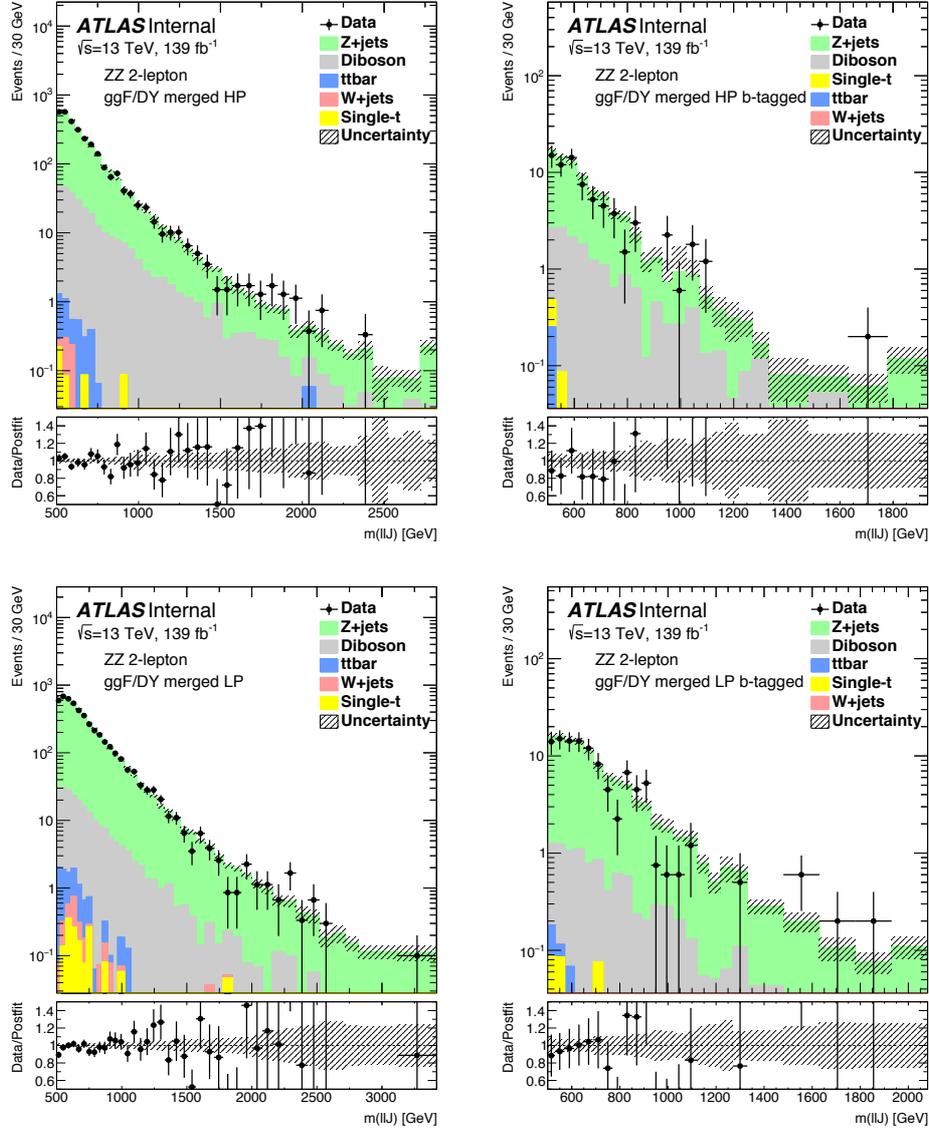


Figure 4.32: Comparisons of the observed data and expected background distributions of the  $m_{ll}$  discriminant in the ggF Merged ZZ SRs. The middle pane shows the ratio of the observed data to the background prediction. The bottom pane show the ratio of post-fit and pre-fit background predictions. The hashed bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.

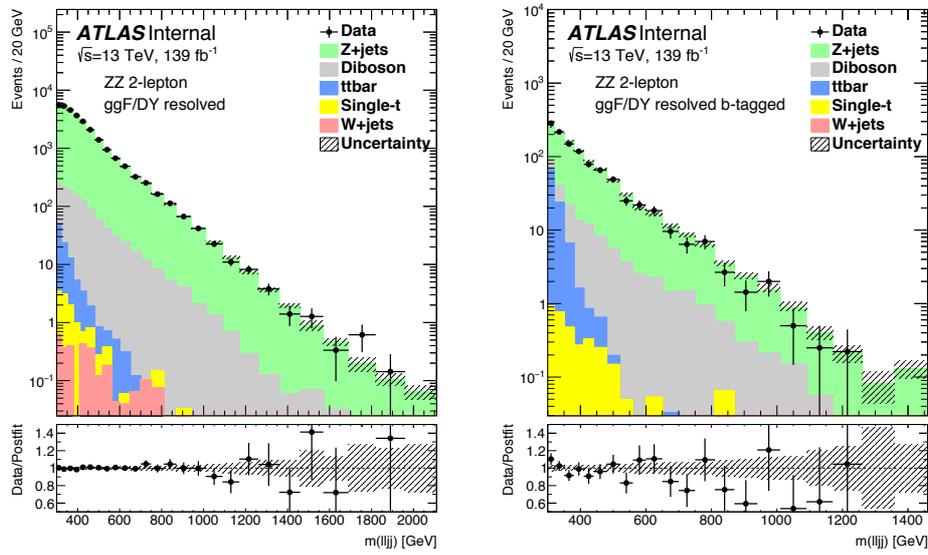


Figure 4.33: Comparisons of the observed data and expected background distributions of the  $m_{Ujj}$  discriminant in the ggF Resolved ZZ SRs. The middle pane shows the ratio of the observed data to the background prediction. The bottom pane show the ratio of post-fit and pre-fit background predictions. The hashed bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.

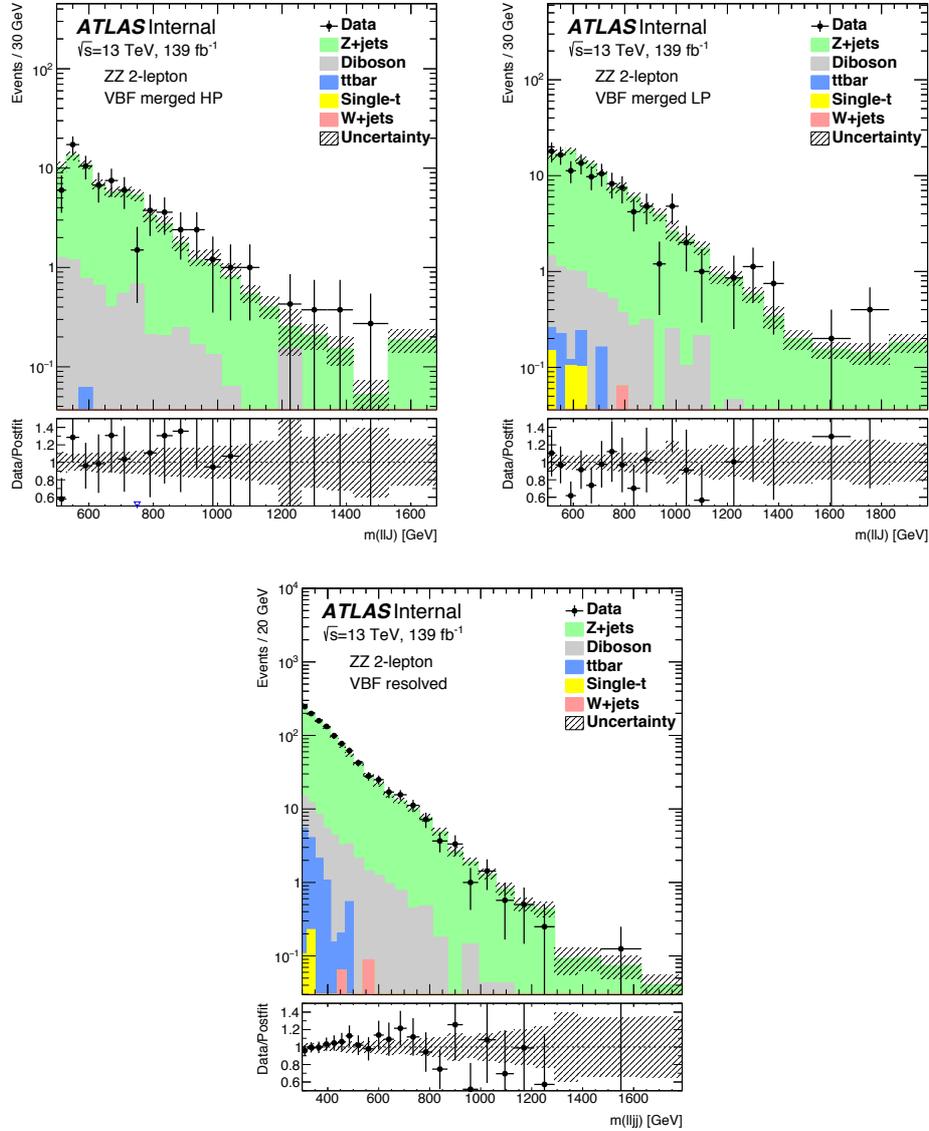


Figure 4.34: Comparisons of the observed data and expected background distributions of the  $m_{ll}$  and  $m_{lljj}$  discriminants in the VBF Merged and Resolved ZZ SRs. The middle pane shows the ratio of the observed data to the background prediction. The bottom pane show the ratio of post-fit and pre-fit background predictions. The hashed bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.

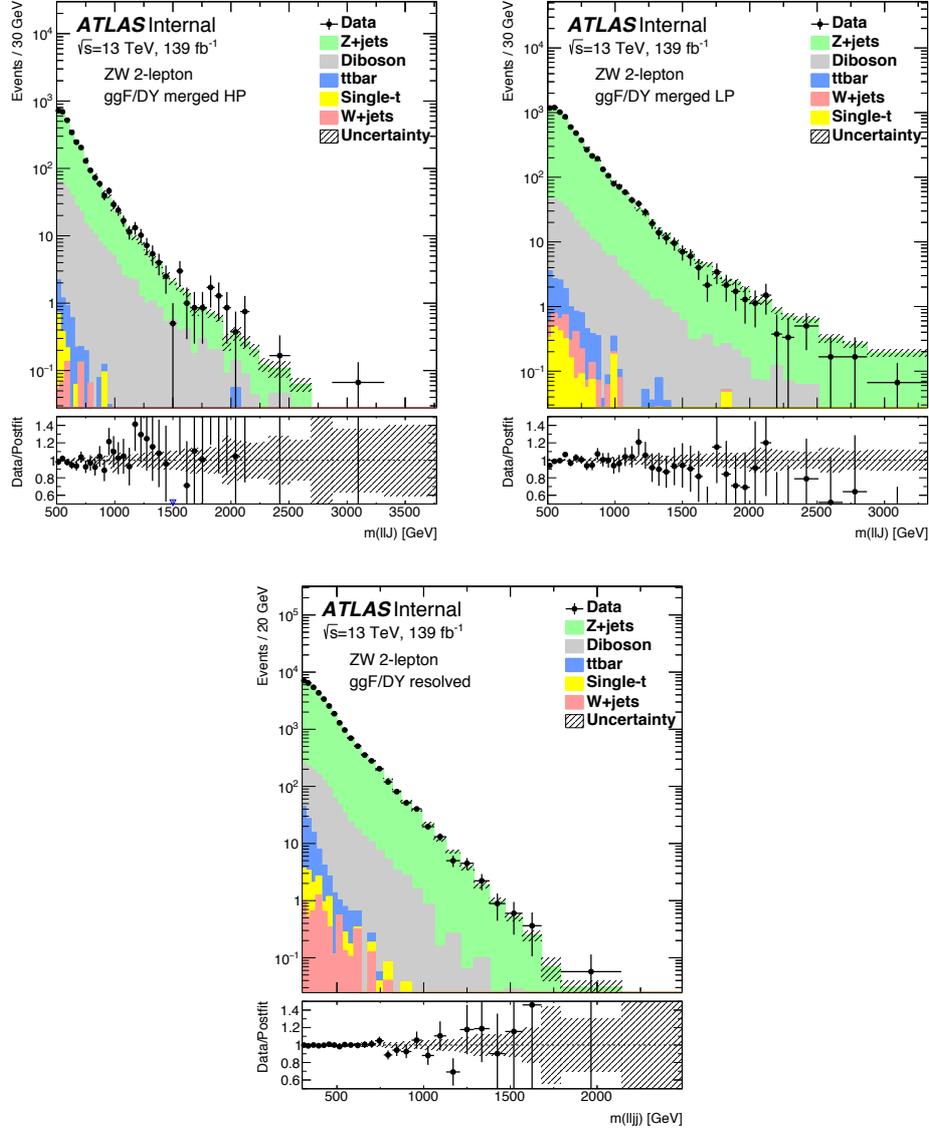


Figure 4.35: Comparisons of the observed data and expected background distributions of the  $m_{lU}$  and  $m_{ljj}$  discriminants in the ggF Merged and Resolved WZ SRs. The middle pane shows the ratio of the observed data to the background prediction. The bottom pane show the ratio of post-fit and pre-fit background predictions. The hashed bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.

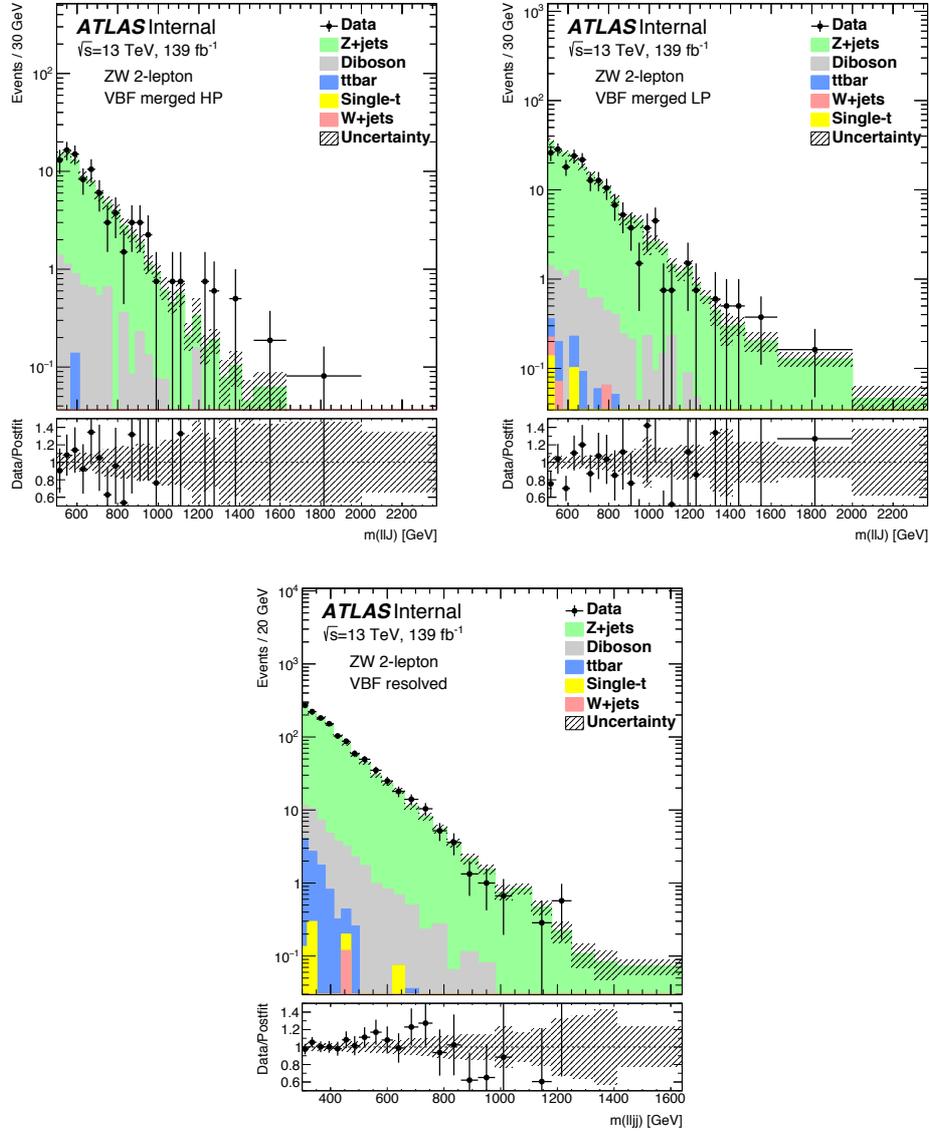


Figure 4.36: Comparisons of the observed data and expected background distributions of the  $m_{ll}$  and  $m_{lljj}$  discriminants in the VBF Merged and Resolved WZ SRs. The middle pane shows the ratio of the observed data to the background prediction. The bottom pane show the ratio of post-fit and pre-fit background predictions. The hashed bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.

pairs are derived by repeating the fit to the signal-plus-background hypothesis for different signal models.

The upper limits on the production cross section of the Radion are obtained by combining searches of  $R \rightarrow WW$  and  $R \rightarrow ZZ$  in the three leptonic final states. The limits are derived, separately for the ggF and VBF processes, through the  $WW + ZZ$  fits for different Radion mass hypotheses. The observed and expected  $\sigma(pp \rightarrow R)$  upper limits at 95% confidence level (CL) as functions of its mass for both the ggF and VBF processes are shown in figure 4.37. The observed limit varies from 1.2 pb at 300 GeV to 0.25 fb at 5 TeV for the ggF production process, and from 0.42 pb at 300 GeV to 0.14 fb at 5 TeV for the VBF production process. Compared with theory cross sections, the observed upper limits exclude the ggF production of a Radion with its mass below 3.4 TeV, while no mass exclusion can be made for the VBF production.

The upper limits on the production cross sections of HVT  $W'$  and  $Z'$  bosons are obtained through the  $WZ$  and  $WW + ZZ$  fits, respectively. All leptonic channels contribute to the  $W' \rightarrow WZ$  search, while only the 1-lepton channel contributes to the  $Z' \rightarrow WW$  search (see tables 4.10 and 4.11). Limits for  $W'$  and  $Z'$  production, as functions of resonance masses, are shown in figures 4.38 and 4.39, for both DY and VBF processes. The observed  $W'$  limit range from 1.9 pb at 300 GeV to 0.12 fb at 5 TeV for DY production, and from 1.3 pb at 300 GeV to 0.3 fb at 4 TeV for VBF production. These limits exclude an HVT  $W'$  boson produced in the DY process lighter than 4.1 TeV for Model A and 4.4 TeV for Model B, but fail to exclude any mass region in the VBF process. Similarly, the observed  $Z'$  limits range from 0.9 pb at 300 GeV to 0.16 fb at 5 TeV for the DY process and from 1.36 pb at 300 GeV to 0.20 fb at 4 TeV for the VBF process. These limits exclude an HVT  $Z'$  boson produced in the DY process lighter than 3.7 TeV for Model A and 4.0 TeV for Model B.

The upper limits on the production cross section of the RS Graviton  $\sigma(pp \rightarrow G_{KK})$  are obtained following the same procedure used to derive the Radion limits. The observed and expected  $\sigma(pp \rightarrow G_{KK})$  upper limits at 95% confidence level, as functions of its mass for both the ggF and VBF processes, are shown in figure 4.40. The observed limit varies from 1.4 pb at 300 GeV to 0.22 fb at 5 TeV for the ggF production process, and from 0.29 pb at 300 GeV to 0.20 fb at 5 TeV for the VBF production process. Compared with theory cross sections, the observed upper limits exclude the production of a RS graviton lighter than 2.0 TeV in the ggF process and lighter than 0.8 TeV in the VBF process.

Channel	$V \rightarrow qq$ recon.	Signal regions	Background estimates							Data
			$W + jets$	$Z + jets$	$t\bar{t}$	$Diboson$	$Single-t$	$Multijet$	Total	
VBF category										
0-lepton (ZZ)	Merged	HP	169 ± 12	228 ± 16	102 ± 10	51 ± 10	24 ± 4	-	(57 ± 3) · 10 <sup>1</sup>	589
		LP	(37 ± 2) · 10 <sup>1</sup>	(41 ± 2) · 10 <sup>1</sup>	75 ± 8	30 ± 4	21 ± 4	-	(91 ± 3) · 10 <sup>1</sup>	936
	Merged	HP Tag	133 ± 14	(27 ± 4) · 10 <sup>1</sup>	(44 ± 3) · 10 <sup>1</sup>	100 ± 10	45 ± 7	-	(98 ± 6) · 10 <sup>1</sup>	978
		LP Tag	(76 ± 4) · 10 <sup>2</sup>	(143 ± 6) · 10 <sup>2</sup>	(60 ± 3) · 10 <sup>2</sup>	(230 ± 19) · 10 <sup>1</sup>	(84 ± 11) · 10 <sup>1</sup>	-	(311 ± 8) · 10 <sup>2</sup>	31074
ggF/DY category										
Merged	HP	(26 ± 3) · 10 <sup>1</sup>	(56 ± 5) · 10 <sup>1</sup>	(34 ± 3) · 10 <sup>1</sup>	67 ± 7	43 ± 7	-	(127 ± 6) · 10 <sup>1</sup>	1277	
	LP	(163 ± 9) · 10 <sup>2</sup>	(286 ± 2) · 10 <sup>2</sup>	(50 ± 2) · 10 <sup>2</sup>	(176 ± 15) · 10 <sup>1</sup>	(60 ± 8) · 10 <sup>1</sup>	-	(524 ± 15) · 10 <sup>2</sup>	52396	
VBF category										
Merged	HP	(53 ± 3) · 10 <sup>1</sup>	8.3 ± 0.5	(32 ± 2) · 10 <sup>1</sup>	(14 ± 3) · 10 <sup>1</sup>	(11 ± 2) · 10 <sup>1</sup>	-	(111 ± 5) · 10 <sup>1</sup>	1096	
	LP	(139 ± 4) · 10 <sup>1</sup>	24.5 ± 1.1	229 ± 17	(15 ± 3) · 10 <sup>1</sup>	83 ± 16	-	(187 ± 6) · 10 <sup>1</sup>	1846	
Resolved		(1136 ± 19) · 10 <sup>1</sup>	526 ± 10	(407 ± 14) · 10 <sup>1</sup>	(59 ± 8) · 10 <sup>1</sup>	(11 ± 2) · 10 <sup>2</sup>	(96 ± 11) · 10 <sup>1</sup>	(186 ± 3) · 10 <sup>2</sup>	18530	
ggF/DY category										
Merged	HP	(2482 ± 17) · 10 <sup>1</sup>	463 ± 5	(139 ± 2) · 10 <sup>1</sup>	(49 ± 3) · 10 <sup>2</sup>	(28 ± 4) · 10 <sup>2</sup>	-	(469 ± 5) · 10 <sup>2</sup>	47330	
	LP	(603 ± 2) · 10 <sup>2</sup>	1096 ± 8	(1105 ± 16) · 10 <sup>1</sup>	(40 ± 2) · 10 <sup>2</sup>	(20 ± 3) · 10 <sup>2</sup>	-	(783 ± 4) · 10 <sup>2</sup>	78380	
Resolved		(4435 ± 18) · 10 <sup>2</sup>	(1248 ± 4) · 10 <sup>1</sup>	(1260 ± 14) · 10 <sup>2</sup>	(168 ± 12) · 10 <sup>2</sup>	(21 ± 3) · 10 <sup>3</sup>	(272 ± 14) · 10 <sup>2</sup>	(647 ± 4) · 10 <sup>3</sup>	645610	
VBF category										
Merged	HP	0	87 ± 6	0.08 ± 0.01	9.6 ± 1.2	0	-	97 ± 6	101	
	LP	0.13 ± 0.01	170 ± 8	0.85 ± 0.07	9.9 ± 1.2	0.43 ± 0.07	-	181 ± 8	162	
Resolved		0.27 ± 0.01	(157 ± 3) · 10 <sup>1</sup>	16.99 ± 0.66	72 ± 10	0.48 ± 0.32	-	(166 ± 3) · 10 <sup>1</sup>	1685	
ggF/DY category										
2-lepton (ZZ)	Merged	HP Tag	0	85 ± 6	0.28 ± 0.03	21 ± 2	0.34 ± 0.05	-	107 ± 7	94
		LP Tag	0.77 ± 0.01	(330 ± 4) · 10 <sup>1</sup>	4.27 ± 0.08	(36 ± 3) · 10 <sup>1</sup>	0.58 ± 0.11	-	(367 ± 5) · 10 <sup>1</sup>	3671
	Merged	HP Untag	0	138 ± 8	0.31 ± 0.03	12.8 ± 1.4	0.30 ± 0.04	-	152 ± 8	141
		LP Untag	2.34 ± 0.02	(592 ± 5) · 10 <sup>1</sup>	10.16 ± 0.16	(28 ± 3) · 10 <sup>1</sup>	2.03 ± 0.29	-	(622 ± 6) · 10 <sup>1</sup>	6095
Resolved	Tag	0	(132 ± 3) · 10 <sup>1</sup>	114 ± 10	159 ± 12	4.74 ± 0.79	-	(160 ± 3) · 10 <sup>1</sup>	1583	
	Untag	4.68 ± 0.03	(4275 ± 16) · 10 <sup>1</sup>	110.6 ± 1.5	(177 ± 10) · 10 <sup>1</sup>	13 ± 2	-	(4465 ± 19) · 10 <sup>1</sup>	44604	

Table 4.15: The numbers of observed data events in all  $WW$  and  $ZZ$  SRs, compared with the expected background events. The backgrounds are estimated from a simultaneous background-only fit to the  $WW$  and  $ZZ$  SRs and their corresponding CRs.

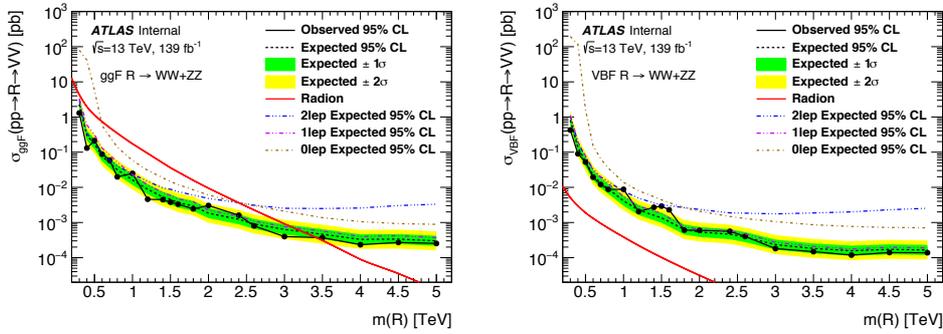


Figure 4.37: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at  $\sqrt{s} = 13$  TeV of a Radion as functions of its mass for the (left plot) ggF and (right plot) VBF processes, combining searches in the three leptonic channels for both the  $WW$  and  $ZZ$  decay modes of the Radion. The green (inner) and yellow (outer) bands represent  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainty in the expected limits. Limits expected from individual leptonic channels (dashed curves in blue, magenta and brown) are also shown for comparison. The theoretical cross sections, as a functions of the Radion mass, are overlaid.

Channel	$V \rightarrow qq$ recon.	Signal regions	Background estimates						Data	
			$W + jets$	$Z + jets$	$t\bar{t}$	Diboson	Single- $t$	Multijet		Total
VBF category										
0-lepton	Merged	HP	$175 \pm 19$	$255 \pm 19$	$(12 \pm 3) \cdot 10^1$	$40 \pm 10$	$23 \pm 5$	-	$(62 \pm 4) \cdot 10^1$	629
		LP	$(48 \pm 4) \cdot 10^1$	$(60 \pm 3) \cdot 10^1$	$78 \pm 18$	$40 \pm 9$	$19 \pm 4$	-	$(121 \pm 6) \cdot 10^1$	1252
	Merged	HP	$(81 \pm 7) \cdot 10^2$	$(153 \pm 8) \cdot 10^2$	$(81 \pm 5) \cdot 10^2$	$(28 \pm 3) \cdot 10^2$	$(104 \pm 16) \cdot 10^1$	-	$(354 \pm 12) \cdot 10^2$	35426
		LP	$(219 \pm 18) \cdot 10^2$	$(39 \pm 2) \cdot 10^3$	$(73 \pm 5) \cdot 10^2$	$(26 \pm 3) \cdot 10^2$	$(86 \pm 13) \cdot 10^1$	-	$(72 \pm 3) \cdot 10^3$	71892
ggF/DY category										
1-lepton	Merged	HP	$(48 \pm 3) \cdot 10^1$	$7.66 \pm 0.51$	$(34 \pm 3) \cdot 10^1$	$(11 \pm 4) \cdot 10^1$	$78 \pm 17$	-	$(102 \pm 6) \cdot 10^1$	1018
		LP	$(99 \pm 4) \cdot 10^1$	$20.76 \pm 0.94$	$186 \pm 18$	$(11 \pm 4) \cdot 10^1$	$59 \pm 13$	-	$(137 \pm 6) \cdot 10^1$	1313
	Resolved	HP	$(102 \pm 3) \cdot 10^2$	$529 \pm 11$	$(50 \pm 2) \cdot 10^2$	$(38 \pm 15) \cdot 10^1$	$(11 \pm 2) \cdot 10^2$	$(77 \pm 13) \cdot 10^1$	$(179 \pm 4) \cdot 10^2$	17826
		LP	$(102 \pm 3) \cdot 10^2$	$529 \pm 11$	$(50 \pm 2) \cdot 10^2$	$(38 \pm 15) \cdot 10^1$	$(11 \pm 2) \cdot 10^2$	$(77 \pm 13) \cdot 10^1$	$(179 \pm 4) \cdot 10^2$	17826
ggF/DY category										
2-lepton	Merged	HP Tag	$(33 \pm 4) \cdot 10^1$	$7.81 \pm 0.80$	$(102 \pm 3) \cdot 10^1$	$100 \pm 11$	$(17 \pm 3) \cdot 10^1$	-	$(163 \pm 6) \cdot 10^1$	1699
		Untag	$(2279 \pm 19) \cdot 10^1$	$428 \pm 5$	$(117 \pm 2) \cdot 10^2$	$(30 \pm 3) \cdot 10^2$	$(21 \pm 3) \cdot 10^2$	-	$(400 \pm 5) \cdot 10^2$	40193
	Merged	LP Tag	$(58 \pm 4) \cdot 10^1$	$11.6 \pm 1.2$	$(71 \pm 3) \cdot 10^1$	$64 \pm 8$	$(15 \pm 3) \cdot 10^1$	-	$(151 \pm 6) \cdot 10^1$	1559
		Untag	$(418 \pm 2) \cdot 10^2$	$766 \pm 6$	$(892 \pm 13) \cdot 10^1$	$(22 \pm 2) \cdot 10^2$	$(14 \pm 2) \cdot 10^2$	-	$(551 \pm 4) \cdot 10^2$	54735
	Resolved	Tag	$(62 \pm 4) \cdot 10^2$	$183 \pm 19$	$(382 \pm 4) \cdot 10^2$	$(62 \pm 6) \cdot 10^1$	$(34 \pm 6) \cdot 10^2$	$(60 \pm 13) \cdot 10^1$	$(492 \pm 8) \cdot 10^2$	48919
		Untag	$(342 \pm 2) \cdot 10^3$	$(996 \pm 3) \cdot 10^1$	$(113 \pm 2) \cdot 10^3$	$(144 \pm 12) \cdot 10^2$	$(18 \pm 3) \cdot 10^3$	$(26 \pm 2) \cdot 10^3$	$(523 \pm 5) \cdot 10^3$	521813
	VBF category									
	2-lepton	Merged	HP	0	$102 \pm 7$	$0.22 \pm 0.02$	$8.9 \pm 1.9$	0	-	$111 \pm 7$
LP			$0.31 \pm 0.02$	$247 \pm 11$	$0.76 \pm 0.07$	$11 \pm 2$	$0.27 \pm 0.06$	-	$259 \pm 11$	243
Resolved		HP	$0.19 \pm 0.01$	$(171 \pm 3) \cdot 10^1$	$12.69 \pm 0.58$	$57 \pm 10$	$0.84 \pm 0.30$	-	$(178 \pm 3) \cdot 10^1$	1831
ggF/DY category										
2-lepton	Merged	HP	$0.73 \pm 0.01$	$(379 \pm 5) \cdot 10^1$	$5.17 \pm 0.10$	$(45 \pm 4) \cdot 10^1$	$1.38 \pm 0.23$	-	$(424 \pm 7) \cdot 10^1$	4197
		LP	$2.98 \pm 0.02$	$(875 \pm 7) \cdot 10^1$	$13.57 \pm 0.21$	$(39 \pm 4) \cdot 10^1$	$2.89 \pm 0.52$	-	$(916 \pm 8) \cdot 10^1$	9088
	Resolved	HP	$8.57 \pm 0.06$	$(5000 \pm 18) \cdot 10^1$	$116 \pm 2$	$(171 \pm 13) \cdot 10^1$	$14 \pm 2$	-	$(519 \pm 2) \cdot 10^2$	51655

Table 4.16: The numbers of observed data events in all  $WZ$  SRs, compared with the expected background events. The backgrounds are estimated from a simultaneous background-only fit to the  $WZ$  SRs and their corresponding CRs.

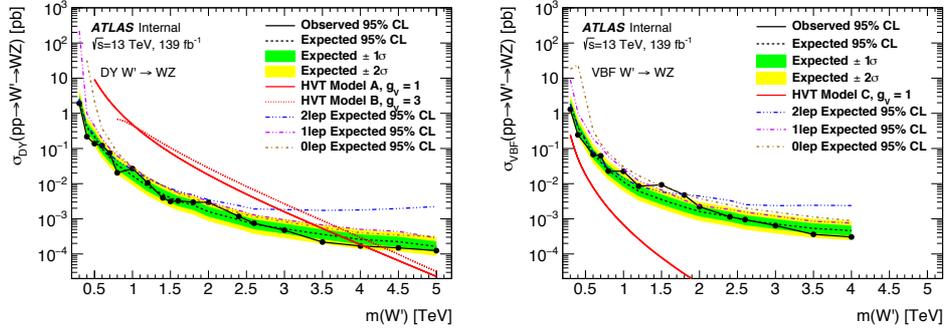


Figure 4.38: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at  $\sqrt{s} = 13$  TeV of HVT  $W'$  as functions of its mass for the (left plot) ggF and (right plot) VBF processes, combining searches in the three leptonic channels for the  $WZ$  decay mode of the  $W'$  resonance. The green (inner) and yellow (outer) bands represent  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainty in the expected limits. Limits expected from individual leptonic channels (dashed curves in blue, magenta and brown) are also shown for comparison. The theoretical cross sections, in Model A, B (left plot) and in Model C (right plot) as a functions of the  $W'$  mass, are overlaid.

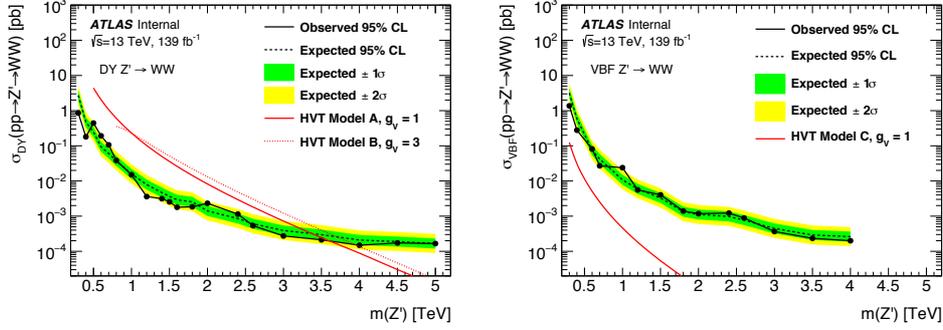


Figure 4.39: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at  $\sqrt{s} = 13$  TeV of HVT  $Z'$  as functions of its mass for the (left plot) ggF and (right plot) VBF processes, for the  $WW$  decay mode of the  $Z'$  resonance. The green (inner) and yellow (outer) bands represent  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainty in the expected limits. The theoretical cross sections, in Model A, B (left plot) and in Model C (right plot) as a functions of the  $Z'$  mass, are overlaid.

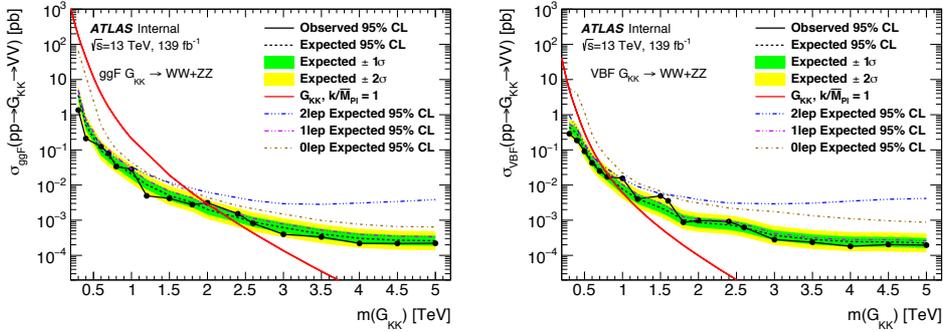


Figure 4.40: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at  $\sqrt{s} = 13$  TeV of a RS Graviton as functions of its mass for the (left plot) ggF and (right plot) VBF processes, combining searches in the three leptonic channels for both the  $WW$  and  $ZZ$  decay modes of the Graviton. The green (inner) and yellow (outer) bands represent  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainty in the expected limits. Limits expected from individual leptonic channels (dashed curves in blue, magenta and brown) are also shown for comparison. The theoretical cross sections, as a functions of the Graviton mass, are overlaid.

### Impact of systematic uncertainties

The effects of systematic uncertainties are studied for hypothesized signals, by evaluating the relative uncertainties in the best-fit signal-strength parameter  $\mu$  from the leading sources of systematic uncertainties, as showed in figure 4.41 for Graviton signal with  $m(X) = 600$  GeV and 2 TeV, respectively. Except for the statistical uncertainties in the data, the uncertainties with the largest impact on the sensitivity of the searches come from the size of the MC samples, floating background normalizations, measurements of small- $R$  and large- $R$  jets, and background modelling. For signals with higher mass, the data statistical uncertainty becomes dominant. The effects of systematic uncertainties for the other signal hypotheses are similar to those shown for the ggF Graviton search.

Tables 4.17 summarizes the relative uncertainties in the best-fit signal-strength parameter  $\mu$ .

$m(G_{kk}) = 600$ GeV		$m(G_{kk}) = 2$ TeV	
Uncertainty source	$\Delta\mu/\mu$ [%]	Uncertainty source	$\Delta\mu/\mu$ [%]
Total systematics	50	Total systematics	59
Pseudo-data statistics	29	Pseudo-data statistics	48
Large- $R$ jet	18	Large- $R$ jet	24
MC statistics	16	MC statistics	17
Floating background normalisations	15	$W/Z$ +jets modelling	15
Diboson modelling	12	Flavor tagging	5.5
$W/Z$ +jets modelling	11	modelling	4.2
Small- $R$ jet	9.7	Diboson modelling	3.9
modelling	8.1	Single- $t$ modelling	3.3

Table 4.17: The relative uncertainties from the leading uncertainty sources in the best-fit signal-strength parameter  $\mu$  of hypothesised signal production of ggF Graviton with  $m(G_{kk}) = 600$  GeV and  $m(G_{kk}) = 2$  TeV. For this study, the Graviton production cross section is assumed to be 100 fb at 600 GeV and 2 fb at 2 TeV, corresponding to approximately the expected median upper limits at these two mass values.

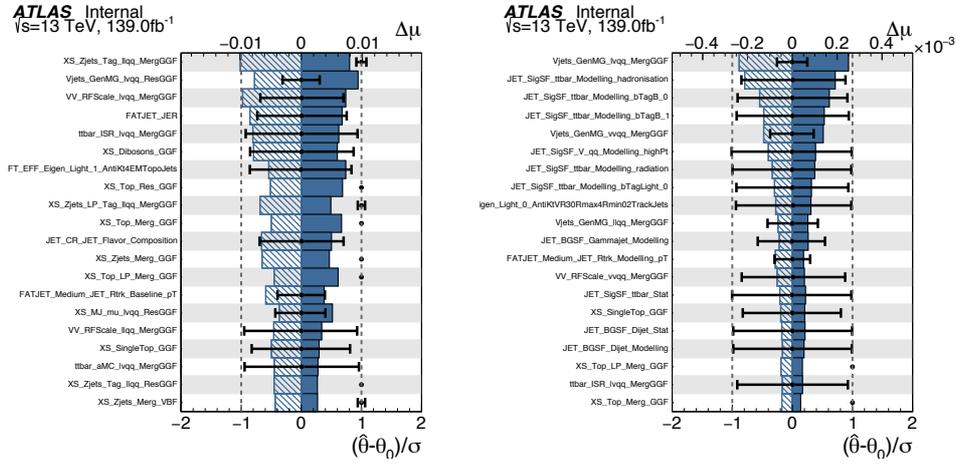


Figure 4.41: Ranking of the nuisance parameter impact on the expected exclusion limits on ggF Graviton  $\sigma(pp \rightarrow X) \times BR(X \rightarrow VV)$ . The Graviton signals with: (left)  $m(G_{kk}) = 600$  GeV and  $\sigma(pp \rightarrow X) \times BR(X \rightarrow VV) = 100$  fb and (right)  $m(G_{kk}) = 2$  TeV and  $\sigma(pp \rightarrow X) \times BR(X \rightarrow VV) = 2$  fb are used for the fit.

# Conclusions

Many models of Physics beyond the Standard Model (SM) predict the existence of a new resonance that could decay in a pair of SM vector bosons. These include extensions to the SM scalar sector as in the two-Higgs-doublet model (2HDM) [20] that predict new spin-0 resonances. In the Heavy Vector Triplet (HVT) phenomenological Lagrangian model [21], a new heavy vector triplet ( $W'$ ,  $Z'$ ) is introduced with the new gauge bosons degenerate in mass. Warped extra dimensions Randall–Sundrum (RS) models [22] predict spin-2 Kaluza–Klein (KK) excitations of the graviton,  $G_{KK}$  [23] and spin-0 Radions [24, 30].

Search for heavy diboson resonances, decaying into a pair of vectorial bosons, has been described in this thesis work. The semileptonic decay mode, in which one boson decays hadronically and the other leptonically, has been investigated. Three final states are possible, depending on the leptonic decay mode:  $X \rightarrow ZV \rightarrow llqq$ ,  $X \rightarrow ZV \rightarrow \nu\nu qq$  and  $X \rightarrow WV \rightarrow lvqq$ . Previous publications with  $36.1 \text{ fb}^{-1}$  data collected in the first part of the Run-II (2015-2016) were performed considering the  $lvqq$  final state alone [82], and a second one combining the  $llqq$  and  $\nu\nu qq$  final states [83].

The work described in this thesis refers mainly to the  $llqq$  final state. As this study is part of a more general analysis that combines  $llqq$ ,  $lvqq$  and  $\nu\nu qq$  final states, the results and, in particular, the upper limits on the cross sections are obtained by combining all these three channels. The dataset used in this analysis is the full available Run-2 statistics collected by the ATLAS experiment at the LHC at  $\sqrt{s} = 13 \text{ TeV}$ . The proton-proton collisions have been recorded during the years 2015-2018, corresponding to an integrated luminosity of  $139 \text{ fb}^{-1}$ .

The hadronic decays of the vector bosons can be identified either as two separate jets (*resolved regime*) or as one large-radius jet (*merged regime*). As the resonance mass increases, the two jets become more collimated and they can be more efficiently reconstructed as one large-radius jet.

My work has been focused on different analysis topics, one of them has been a dedicated study to suppress objects that are falsely reconstructed as prompt leptons, e.g. from semileptonic heavy-flavour hadron decays or misidentified jets. The leptons in the event are required to be isolated, investigating the total deposited transverse energy of topological clusters and the sum of the transverse momenta of all tracks in a region around the identified electron; calorimeter isolation, track isolation or both criteria can be applied and I performed a detailed investigation of the isolation requirements for the leptons selected in the  $llqq$  analysis. Most of my work has also been dedicated on the optimization of the analysis events selection, in order to enhance the sensitivity to heavy diboson resonance searches.

The main background processes for the  $llqq$  analysis are the  $Z$ +jets, top quark and SM diboson production. Heavy resonances would manifest themselves as resonant structures above the SM background in the invariant-mass distribution of the final state.

Upper limits on the production cross sections of heavy resonances times their decay branching ratios to the diboson couple are derived in the mass range 300-5000 GeV, within the context of Beyond Standard Model theories with an heavy vector triplet, or Gravitons and Radions in extra dimensions. Production through gluon-gluon fusion, Drell-Yan or vector-boson fusion are considered, depending on the assumed model.

A combined profile likelihood fit, to binned discriminants in all categories and regions simultaneously, has been performed. The final discriminants are the invariant masses in the resolved and merged final states. The test statistic, constructed according to the profile likelihood, is used to measure the compatibility of the background-only model with the observed data and for exclusion intervals derived with the  $CL_s$  method [117, 118]. The limits are set on the signal cross section times branching ratio, using the theoretical cross section and branching ratio for the given signal model.

The limits have been derived combining the three lepton final states:  $llqq$ ,  $lvqq$  and  $\nu\nu qq$ . The data are reasonably well reproduced by the background contributions in all the final discriminant distributions and no significant excess have been observed in any of the three lepton channels. In absence of significant deviations from expected background contributions, constraints on the production cross sections of heavy resonances decaying to  $VV$  pairs are derived by repeating the fit to the signal-plus-background hypothesis for different signal models.

Upper limits on the production cross section of the Radion are obtained separately for the ggF and VBF processes. The observed limit varies from 1.2 pb at 300 GeV to 0.25 fb at 5 TeV for the ggF production process, and from

0.42 pb at 300 GeV to 0.14 fb at 5 TeV for the VBF production process. Compared with theory cross sections, the observed upper limits exclude the ggF production of a Radion with its mass below 3.4 TeV, while no mass exclusion can be made for the VBF production.

Upper limits on the production cross sections of HVT  $W'$  and  $Z'$  bosons are obtained. The observed  $W'$  limit range from 1.9 pb at 300 GeV to 0.12 fb at 5 TeV for DY production, and from 1.3 pb at 300 GeV to 0.3 fb at 4 TeV for VBF production. These limits exclude an HVT  $W'$  boson produced in the DY process lighter than 4.1 TeV for Model A and 4.4 TeV for Model B, but fail to exclude any mass region in the VBF process. Similarly, the observed  $Z'$  limits range from 0.9 pb at 300 GeV to 0.16 fb at 5 TeV for the DY process and from 1.36 pb at 300 GeV to 0.20 fb at 4 TeV for the VBF process. These limits exclude an HVT  $Z'$  boson produced in the DY process lighter than 3.7 TeV for Model A and 4.0 TeV for Model B.

Even if the VBF process can't be excluded in these models, the limits on cross sections are an important result since they can be reinterpreted in other models characterized by the same experimental signature of the benchmark models studied.

Upper limits on the production cross section of the RS Graviton are also obtained. The observed limit varies from 3.6 pb at 300 GeV to 0.26 fb at 5 TeV for the ggF production process, and from 0.43 pb at 300 GeV to 0.22 fb at 5 TeV for the VBF production process. Compared with theory cross sections, the observed upper limits exclude the production of a RS Graviton lighter than 1.88 TeV in the ggF process and lighter than 0.84 TeV in the VBF process.

Several improvements in the analysis have been introduced with respect to the previous results, and the dataset analysed is approximately 4 times larger than the previous one. This has led to an improvement in the expected limits for all the signal interpretations and the results will be published in paper before the summer 2020.

Despite the excellent results obtained, further improvements in the  $llqq$  channel are possible. As it possible to see from the limit plots showed in section 4.6.2, the limits in the high resonance mass region for the  $llqq$  channel are worse than the ones in the  $l\nu qq$  and  $\nu\nu qq$  research channels. As for the hadronic decay, also the leptons from boosted bosons are close together. So, for high mass resonance the two leptons result not isolated, bringing in an efficiency loss and a consequent loss of the analysis sensitivity. Improvements in the identification and isolation of the leptons in the boosted regime can lead to a big improvement of the  $llqq$  channel.

Moreover, a machine learning approach has proven to be effective for a good event classification based on the resonance production mechanism, and in principle this approach can be used for other analysis topic, like the signal to background classification or the discrimination of the signal merged jets from the ones coming from the QCD background.

Great results have been derived and further improvements are still possible in the search of heavy resonances in the  $llqq$  channel.

# Appendix A

## Standard Model production cross sections

Figure A.1 reports several Standard Model total production cross section measurements, corrected for branching fractions, and compared to the corresponding theoretical expectations.

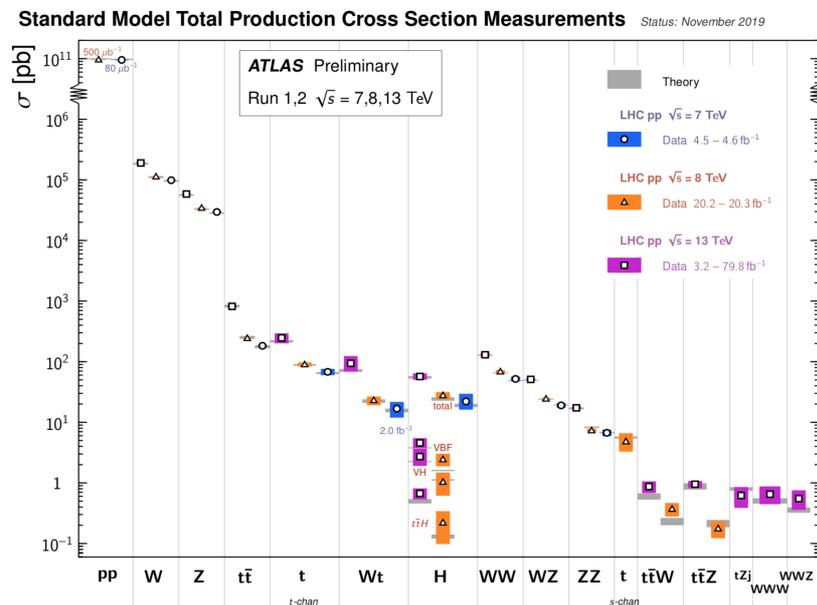


Figure A.1: Standard Model total production cross section measurements, corrected for branching fractions, and compared to the corresponding theoretical expectations [119].

Tables A.1, A.2, A.3, A.4, A.5, A.6, A.7, A.8 summarize the production cross section times branching ratios for the  $Z + jets$ ,  $W + jets$ , top and diboson components.

Table A.1:  $Z \rightarrow ee+jets$  production cross sections, divided per  $p_T$  slices.

$p_T$ slice	flavor component	$\sigma \times \text{BR}$ [pb]
0-70 GeV	Zee + light flavor jets	1.982E+00
	Zee + c-flavor jets	1.981E+00
	Zee + b-flavor jets	1.982E+00
70-140 GeV	Zee + light flavor jets	1.105E-01
	Zee + c-flavor jets	1.106E-01
	Zee + b-flavor jets	1.103E-01
140-280 GeV	Zee + light flavor jets	4.073E-02
	Zee + c-flavor jets	4.067E-02
	Zee + b-flavor jets	4.069E-02
280-500 GeV	Zee + light flavor jets	8.674E-03
	Zee + c-flavor jets	8.671E-03
	Zee + b-flavor jets	8.677E-03
500-1000 GeV	Zee+jets (inclusive)	1.808E-03
> 1000 GeV	Zee+jets (inclusive)	1.486E-04

Table A.2:  $Z \rightarrow \mu\mu$ +jets production cross sections, divided per  $p_T$  slices.

$p_T$ slice	flavor component	$\sigma \times \text{BR}$ [pb]
0-70 GeV	$Z\mu\mu$ + light flavor jets	1.983E+00
	$Z\mu\mu$ + c-flavor jets	1.978E+00
	$Z\mu\mu$ + b-flavor jets	1.982E+00
70-140 GeV	$Z\mu\mu$ + light flavor jets	1.089E-01
	$Z\mu\mu$ + c-flavor jets	1.094E-01
	$Z\mu\mu$ + b-flavor jets	1.094E-01
140-280 GeV	$Z\mu\mu$ + light flavor jets	3.988E-02
	$Z\mu\mu$ + c-flavor jets	3.980E-02
	$Z\mu\mu$ + b-flavor jets	3.991E-02
280-500 GeV	$Z\mu\mu$ + light flavor jets	8.537E-03
	$Z\mu\mu$ + c-flavor jets	8.540E-03
	$Z\mu\mu$ + b-flavor jets	8.493E-03
500-1000 GeV	$Z\mu\mu$ +jets (inclusive)	1.788E-03
> 1000 GeV	$Z\mu\mu$ +jets (inclusive)	1.477E-04

Table A.3:  $Z \rightarrow \tau\tau$ +jets production cross sections, divided per  $p_T$  slices.

$p_T$ slice	flavor component	$\sigma \times \text{BR}$ [pb]
0-70 GeV	$Z\tau\tau$ + light flavor jets	1.982E+00
	$Z\tau\tau$ + c-flavor jets	1.979E+00
	$Z\tau\tau$ + b-flavor jets	1.982E+00
70-140 GeV	$Z\tau\tau$ + light flavor jets	1.104E-01
	$Z\tau\tau$ + c-flavor jets	1.105E-01
	$Z\tau\tau$ + b-flavor jets	1.109E-01
140-280 GeV	$Z\tau\tau$ + light flavor jets	4.078E-02
	$Z\tau\tau$ + c-flavor jets	4.074E-02
	$Z\tau\tau$ + b-flavor jets	4.076E-02
280-500 GeV	$Z\tau\tau$ + light flavor jets	8.550E-03
	$Z\tau\tau$ + c-flavor jets	8.671E-03
	$Z\tau\tau$ + b-flavor jets	8.680E-03
500-1000 GeV	$Z\tau\tau$ +jets (inclusive)	1.810E-03
> 1000 GeV	$Z\tau\tau$ +jets (inclusive)	1.483E-04

Table A.4:  $W \rightarrow e\nu$ +jets production cross sections, divided per  $p_T$  slices.

$p_T$ slice	flavor component	$\sigma \times \text{BR}$ [pb]
0-70 GeV	$W_{e\nu}$ + light flavor jets	1.913E+01
	$W_{e\nu}$ + c-flavor jets	1.913E+01
	$W_{e\nu}$ + b-flavor jets	1.914E+01
70-140 GeV	$W_{e\nu}$ + light flavor jets	9.426E-01
	$W_{e\nu}$ + c-flavor jets	9.457E-01
	$W_{e\nu}$ + b-flavor jets	9.451E-01
140-280 GeV	$W_{e\nu}$ + light flavor jets	3.398E-01
	$W_{e\nu}$ + c-flavor jets	3.399E-01
	$W_{e\nu}$ + b-flavor jets	3.395E-01
280-500 GeV	$W_{e\nu}$ + light flavor jets	7.208E-02
	$W_{e\nu}$ + c-flavor jets	7.213E-02
	$W_{e\nu}$ + b-flavor jets	7.211E-02
500-1000 GeV	$W_{e\nu}$ +jets (inclusive)	1.522E-02
> 1000 GeV	$W_{e\nu}$ +jets (inclusive)	1.233E-03

Table A.5:  $W \rightarrow \mu\nu$ +jets production cross sections, divided per  $p_T$  slices.

$p_T$ slice	flavor component	$\sigma \times \text{BR}$ [pb]
0-70 GeV	$W_{\mu\nu}$ + light flavor jets	1.914E+01
	$W_{\mu\nu}$ + c-flavor jets	1.912E+01
	$W_{\mu\nu}$ + b-flavor jets	1.914E+01
70-140 GeV	$W_{\mu\nu}$ + light flavor jets	9.448E-01
	$W_{\mu\nu}$ + c-flavor jets	9.378E-01
	$W_{\mu\nu}$ + b-flavor jets	9.446E-01
140-280 GeV	$W_{\mu\nu}$ + light flavor jets	3.395E-01
	$W_{\mu\nu}$ + c-flavor jets	3.401E-01
	$W_{\mu\nu}$ + b-flavor jets	3.395E-01
280-500 GeV	$W_{\mu\nu}$ + light flavor jets	7.207E-02
	$W_{\mu\nu}$ + c-flavor jets	7.220E-02
	$W_{\mu\nu}$ + b-flavor jets	7.204E-02
500-1000 GeV	$W_{\mu\nu}$ +jets (inclusive)	1.501E-02
> 1000 GeV	$W_{\mu\nu}$ +jets (inclusive)	1.234E-03

Table A.6:  $W \rightarrow \tau\nu$ +jets production cross sections, divided per  $p_T$  slices.

$p_T$ slice	flavor component	$\sigma \times \text{BR}$ [pb]
0-70 GeV	$W\tau\nu$ + light flavor jets	1.915E+01
	$W\tau\nu$ + c-flavor jets	1.915E+01
	$W\tau\nu$ + b-flavor jets	1.916E+01
70-140 GeV	$W\tau\nu$ + light flavor jets	9.476E-01
	$W\tau\nu$ + c-flavor jets	9.467E-01
	$W\tau\nu$ + b-flavor jets	9.433E-01
140-280 GeV	$W\tau\nu$ + light flavor jets	3.394E-01
	$W\tau\nu$ + c-flavor jets	3.396E-01
	$W\tau\nu$ + b-flavor jets	3.395E-01
280-500 GeV	$W\tau\nu$ + light flavor jets	7.207E-02
	$W\tau\nu$ + c-flavor jets	7.198E-02
	$W\tau\nu$ + b-flavor jets	7.203E-02
500-1000 GeV	$W\tau\nu$ +jets (inclusive)	1.505E-02
> 1000 GeV	$W\tau\nu$ +jets (inclusive)	1.234E-03

Table A.7: Top production cross sections in dilepton final state.

Sample	$\sigma \times \text{BR}$ [pb]
$t\bar{t}$	8.318E-01
Single top s-channel	1.267E-03
Single top t-channel	2.218E-02
Single top Wt-channel	3.994E-03

Table A.8: Diboson production cross sections.

Sample	$\sigma \times \text{BR}$ [pb]
$Z \rightarrow qqZ \rightarrow ll$	1.556E-02
$W \rightarrow qqZ \rightarrow ll$	3.433E-03
$W \rightarrow qqW \rightarrow l\nu$	2.473E-02
$Z \rightarrow qqW \rightarrow l\nu$	1.142E-02

# Bibliography

- [1] A. Bettini, *Introduction to Elementary Particle Physics*. Cambridge University Press, 2008.
- [2] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: A practical introduction. Vol. 1: From relativistic quantum mechanics to QED*. CRC Press, Bristol, UK, 2012.
- [3] W. E. Burcham and M. Jobes, *Nuclear and particle physics*. Longmans, Essex, 1995. <https://cds.cern.ch/record/285607>.
- [4] D. J. Griffiths, *Introduction to elementary particles*. TextBook Physics. Wiley, New York, NY, 1987. <https://cds.cern.ch/record/1260972>.
- [5] E. Leader and E. Predazzi, *An introduction to gauge theories and the "new physics"*. Cambridge Univ. Press, Cambridge, 1982. <https://cds.cern.ch/record/100166>.
- [6] P. B. Renton, *Electroweak Interactions: An Introduction to the Physics of Quarks and Leptons*. Cambridge Univ. Press, Cambridge, 1990. <https://cds.cern.ch/record/206324>.
- [7] P. A. M. Dirac, *The Quantum Theory of the Electron*, Proceedings of the Royal Society of London Series A **117** no. 778, (1928) 610–624.
- [8] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191–195, [,150(1954)].
- [9] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and collider physics*. Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge University Press, Cambridge, 2003. <https://cds.cern.ch/record/318585>. Photography by S. Vascotto.

- [10] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.
- [11] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266, <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [12] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C680519 (1968-05-19/1968-05-25) 367.
- [13] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. 1984.
- [14] Particle Data Group Collaboration, M. et al., *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001, <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [15] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323, [,157(1964)].
- [16] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex].
- [17] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, [arXiv:1207.7235](https://arxiv.org/abs/1207.7235) [hep-ex].
- [18] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, and et al., *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, The European Physical Journal C **75** no. 10, (2015), <http://dx.doi.org/10.1140/epjc/s10052-015-3685-1>.
- [19] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, R. Aben, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al., *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at  $\sqrt{s} = 7$  and 8 TeV*, Journal of High Energy Physics **2016** no. 8, (2016), [http://dx.doi.org/10.1007/JHEP08\(2016\)045](http://dx.doi.org/10.1007/JHEP08(2016)045).
- [20] G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, and J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, Physics Reports **516** no. 1-2, (2012) 1–102, <http://dx.doi.org/10.1016/j.physrep.2012.02.002>.

- [21] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, *Heavy vector triplets: bridging theory and data*, Journal of High Energy Physics **2014** no. 9, (2014), [http://dx.doi.org/10.1007/JHEP09\(2014\)060](http://dx.doi.org/10.1007/JHEP09(2014)060).
- [22] L. Randall and R. Sundrum, *Large Mass Hierarchy from a Small Extra Dimension*, Physical Review Letters **83** no. 17, (1999) 3370–3373, <http://dx.doi.org/10.1103/PhysRevLett.83.3370>.
- [23] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, *Warped gravitons at the CERN LHC and beyond*, Physical Review D **76** no. 3, (2007), <http://dx.doi.org/10.1103/PhysRevD.76.036006>.
- [24] W. D. Goldberger and M. B. Wise, *Modulus Stabilization with Bulk Fields*, Physical Review Letters **83** no. 24, (1999) 4922–4925, <http://dx.doi.org/10.1103/PhysRevLett.83.4922>.
- [25] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. D **8** (1973) 1226–1239, <https://link.aps.org/doi/10.1103/PhysRevD.8.1226>.
- [26] J. E. Kim, *Light Pseudoscalars, Particle Physics and Cosmology*, Phys. Rept. **150** (1987) 1–177.
- [27] J. M. Cornwall, D. N. Levin, and G. Tiktopoulos, *Erratum: Derivation of gauge invariance from high-energy unitarity bounds on the S matrix*, Phys. Rev. D **11** (1975) 972–972, <https://link.aps.org/doi/10.1103/PhysRevD.11.972>.
- [28] G. J. Gounaris, R. Kögerler, and H. Neufeld, *Relationship between longitudinally polarized vector bosons and their unphysical scalar partners*, Phys. Rev. D **34** (1986) 3257–3259, <https://link.aps.org/doi/10.1103/PhysRevD.34.3257>.
- [29] M. S. Chanowitz and M. K. Gaillard, *The TeV physics of strongly interacting W's and Z's*, Nuclear Physics B **261** (1985) 379 – 431, <http://www.sciencedirect.com/science/article/pii/0550321385905802>.
- [30] W. D. Goldberger and M. B. Wise, *Phenomenology of a stabilized modulus*, Physics Letters B **475** no. 3-4, (2000) 275–279, [http://dx.doi.org/10.1016/S0370-2693\(00\)00099-X](http://dx.doi.org/10.1016/S0370-2693(00)00099-X).
- [31] A. Carvalho, *Gravity particles from Warped Extra Dimensions, predictions for LHC*, 2014.

- [32] V. Barger and M. Ishida, *Randall–Sundrum reality at the LHC*, Physics Letters B **709** no. 3, (2012) 185–191, <http://dx.doi.org/10.1016/j.physletb.2012.01.073>.
- [33] C. Csáki, J. Hubisz, and S. J. Lee, *Radion phenomenology in realistic warped space models*, Physical Review D **76** no. 12, (2007), <http://dx.doi.org/10.1103/PhysRevD.76.125015>.
- [34] V. Barger, P. Langacker, M. McCaskey, M. J. Ramsey-Musolf, and G. Shaughnessy, *CERN LHC phenomenology of an extended standard model with a real scalar singlet*, Physical Review D **77** no. 3, (2008), <http://dx.doi.org/10.1103/PhysRevD.77.035005>.
- [35] C. Oleari, *The POWHEG BOX*, Nuclear Physics B - Proceedings Supplements **205-206** (2010) 36–41, <http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.016>.
- [36] V. Barger, W. Y. Keung, and E. Ma, *Gauge model with light W and Z bosons*, Phys. Rev. D **22** (1980) 727–737, <https://link.aps.org/doi/10.1103/PhysRevD.22.727>.
- [37] R. Contino, D. Pappadopulo, D. Marzocca, and R. Rattazzi, *On the effect of resonances in composite Higgs phenomenology*, Journal of High Energy Physics **2011** no. 10, (2011), [http://dx.doi.org/10.1007/JHEP10\(2011\)081](http://dx.doi.org/10.1007/JHEP10(2011)081).
- [38] The Large Hadron Collider webpage. <http://lhc.web.cern.ch/lhc/>.
- [39] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [40] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [41] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [42] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [43] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [44] W. de Boer, *Precision Experiments at LEP, 60 Years of CERN Experiments and Discoveries* (2015) 107–136, [http://dx.doi.org/10.1142/9789814644150\\_005](http://dx.doi.org/10.1142/9789814644150_005).

- [45] ATLAS Collaboration, *Luminosity determination in pp collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC*, <https://cds.cern.ch/record/2677054>.
- [46] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 1*, tech. rep., 1997.
- [47] ATLAS Collaboration, *ATLAS calorimeter performance Technical Design Report*, tech. rep., 1996.
- [48] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. <https://cds.cern.ch/record/331068>.
- [49] ATLAS Collaboration, *Performance of the ATLAS RPC detector and L1 Muon Barrel trigger at  $\sqrt{s} = 13$  TeV*, <https://cds.cern.ch/record/2668392>.
- [50] ATLAS Collaboration, *ATLAS Trigger and DAQ Home Page*, <http://atlas.web.cern.ch/Atlas/GROUPS/DAQTRIG/daqtrig.html>.
- [51] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, Tech. Rep. ATL-DAQ-PUB-2016-001, CERN, Geneva, Mar, 2016. <https://cds.cern.ch/record/2136007>.
- [52] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter clustering algorithms: Description and performance*,.
- [53] A. Salzburger, *Track Simulation and Reconstruction in the ATLAS experiment*, [http://physik.uibk.ac.at/hephy/theses/diss\\_a\\_s.pdf](http://physik.uibk.ac.at/hephy/theses/diss_a_s.pdf).
- [54] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*,.
- [55] ATLAS Collaboration, *Electron efficiency measurements in 2017 data and electron identification discriminating variables from 2016 data*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2018-002/index.html>.
- [56] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, <https://cds.cern.ch/record/2157687>.

- [57] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. A. Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, and et al., *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data*, The European Physical Journal C **74** no. 7, (2014), <http://dx.doi.org/10.1140/epjc/s10052-014-2941-0>.
- [58] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, The European Physical Journal C **77** no. 3, (2017) 195, <https://doi.org/10.1140/epjc/s10052-017-4756-2>.
- [59] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, <http://cds.cern.ch/record/2157687>.
- [60] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, <http://cds.cern.ch/record/1449796>.
- [61] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. **C76** no. 5, (2016) 292, [arXiv:1603.05598](https://arxiv.org/abs/1603.05598) [hep-ex].
- [62] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, <https://cds.cern.ch/record/1099735>.
- [63] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- $k$   $t$  jet clustering algorithm*, Journal of High Energy Physics **2008** no. 04, (2008) 063, <http://stacks.iop.org/1126-6708/2008/i=04/a=063>.
- [64] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross sections in deep-inelastic scattering*, arXiv preprint hep-ph/9907280, (1999).
- [65] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Physical Review D **48** no. 07, (1993) 3160.
- [66] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-018, CERN, Geneva, May, 2014. <http://cds.cern.ch/record/1700870>.

- [67] 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) 076, arXiv:1306.4945 [hep-ex].
- [68] D. Krohn, J. Thaler, and L.-T. Wang, *Jet trimming*, Journal of High Energy Physics **2010** no. 2, (2010), [http://dx.doi.org/10.1007/JHEP02\(2010\)084](http://dx.doi.org/10.1007/JHEP02(2010)084).
- [69] ATLAS Collaboration, *Improving jet substructure performance in ATLAS using Track-CaloClusters*, Tech. Rep. ATL-PHYS-PUB-2017-015, CERN, Geneva, Jul, 2017. <https://cds.cern.ch/record/2275636>.
- [70] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at  $\sqrt{s} = 13$  TeV*, Tech. Rep. ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul, 2015. <http://cds.cern.ch/record/2037613>.
- [71] M. Aharrouche, H. Ma, C. Adam-Bourdarios, M. Aleksa, D. Banfi, D. Benckroun, K. Benslama, M. Boonekamp, T. Carli, L. Carminati, H. Chen, M. Citterio, D. Dannheim, M. Delmastro, F. Derue, B. Di Girolamo, M. El Kacimi, M. Fanti, R. Froeschl, and I. Wingerter-Seez, *Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam*, Nuclear Instruments and Methods in Physics Research A **614** (2010) .
- [72] E. Abat et al., *Combined performance studies for electrons at the 2004 ATLAS combined test-beam*, JINST **5** (2010) P11006.
- [73] ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*,.
- [74] G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abidinov, B. Abi, M. Abolins, H. Abramowicz, and et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV*, The European Physical Journal C **73** no. 3, (2013), <http://dx.doi.org/10.1140/epjc/s10052-013-2304-2>.
- [75] ATLAS Collaboration, *Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at  $\sqrt{s} = 900$  GeV*,.

- [76] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** no. 3, (2013) 2306, [arXiv:1210.6210](https://arxiv.org/abs/1210.6210) [hep-ex].
- [77] ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, 2012.
- [78] A. L. Rosa, *The ATLAS Insertable B-Layer: from construction to operation*, 2016.
- [79] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016. <http://cds.cern.ch/record/2160731>.
- [80] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at  $\sqrt{s} = 13$  TeV*, Tech. Rep. ATL-PHYS-PUB-2015-023, CERN, Geneva, Jul, 2015. <http://cds.cern.ch/record/2037700>.
- [81] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at  $\sqrt{s} = 13$  TeV*, Tech. Rep. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015. <http://cds.cern.ch/record/2037904>.
- [82] ATLAS Collaboration, *Search for WW/WZ resonance production in  $lvqq$  final states in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Journal of High Energy Physics **2018** no. 3, (2018), [http://dx.doi.org/10.1007/jhep03\(2018\)042](http://dx.doi.org/10.1007/jhep03(2018)042).
- [83] ATLAS Collaboration, *Searches for heavy ZZ and ZW resonances in the  $llqq$  and  $\nu\nu qq$  final states in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Journal of High Energy Physics **2018** no. 3, (2018), [http://dx.doi.org/10.1007/jhep03\(2018\)009](http://dx.doi.org/10.1007/jhep03(2018)009).
- [84] ATLAS Collaboration, *Luminosity determination in pp collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC*, Tech. Rep. ATLAS-CONF-2019-021, CERN, Geneva, Jun, 2019. <https://cds.cern.ch/record/2677054>.
- [85] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** no. 07, (2018) P07017.

- [86] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, The European Physical Journal C **70** no. 3, (2010) 823–874, <https://doi.org/10.1140/epjc/s10052-010-1429-9>.
- [87] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.
- [88] T. Sjöstrand, *PYTHIA 8 Status Report*, 2008.
- [89] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155.
- [90] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, Journal of High Energy Physics **2009** no. 02, (2009) 007–007, <http://dx.doi.org/10.1088/1126-6708/2009/02/007>.
- [91] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, Journal of High Energy Physics **2008** no. 12, (2008) 039–039, <http://dx.doi.org/10.1088/1126-6708/2008/12/039>.
- [92] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Physical Review Letters **108** no. 11, (2012), <http://dx.doi.org/10.1103/PhysRevLett.108.111601>.
- [93] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, Journal of High Energy Physics **2008** no. 03, (2008) 038–038, <http://dx.doi.org/10.1088/1126-6708/2008/03/038>.
- [94] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, Journal of High Energy Physics **2013** no. 4, (2013), [http://dx.doi.org/10.1007/JHEP04\(2013\)027](http://dx.doi.org/10.1007/JHEP04(2013)027).
- [95] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, and et al., *Parton distributions for the LHC run II*, Journal of High Energy Physics **2015** no. 4, (2015), [http://dx.doi.org/10.1007/JHEP04\(2015\)040](http://dx.doi.org/10.1007/JHEP04(2015)040).
- [96] C. Anastasiou, L. Dixon, K. Melnikov, and F. Petriello, *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, Phys. Rev. D **69** (2004) 094008, <https://link.aps.org/doi/10.1103/PhysRevD.69.094008>.

- [97] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics **2010** no. 6, (2010), [http://dx.doi.org/10.1007/JHEP06\(2010\)043](http://dx.doi.org/10.1007/JHEP06(2010)043).
- [98] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, Journal of High Energy Physics **2013** no. 3, (2013), [http://dx.doi.org/10.1007/JHEP03\(2013\)015](http://dx.doi.org/10.1007/JHEP03(2013)015).
- [99] A. Buckley and D. Kar, *ATLAS Run 1 Pythia8 tunes*, Tech. Rep. ATL-COM-PHYS-2014-1337, CERN, Geneva, Oct, 2014. <https://cds.cern.ch/record/1954847>. To be ATL-PHYS-PUB-2014-021.
- [100] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Computer Physics Communications **185** no. 11, (2014) 2930–2938, <http://dx.doi.org/10.1016/j.cpc.2014.06.021>.
- [101] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for channel single top quark production*, Physical Review D **83** no. 9, (2011), <http://dx.doi.org/10.1103/PhysRevD.83.091503>.
- [102] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for channel single top quark production*, Physical Review D **81** no. 5, (2010), <http://dx.doi.org/10.1103/PhysRevD.81.054028>.
- [103] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a  $W$  or  $Z$* , Physical Review D **82** no. 5, (2010), <http://dx.doi.org/10.1103/PhysRevD.82.054018>.
- [104] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, Journal of High Energy Physics **2014** no. 7, (2014), [http://dx.doi.org/10.1007/JHEP07\(2014\)079](http://dx.doi.org/10.1007/JHEP07(2014)079).
- [105] A. J. Larkoski, I. Moult, and D. Neill, *Power counting to better jet observables*, Journal of High Energy Physics **2014** no. 12, (2014), [http://dx.doi.org/10.1007/JHEP12\(2014\)009](http://dx.doi.org/10.1007/JHEP12(2014)009).

- [106] A. J. Larkoski, I. Moutl, and D. Neill, *Analytic boosted boson discrimination*, Journal of High Energy Physics **2016** no. 5, (2016), [http://dx.doi.org/10.1007/JHEP05\(2016\)117](http://dx.doi.org/10.1007/JHEP05(2016)117).
- [107] ATLAS Collaboration, *Search for diboson resonances in hadronic final states in 139 fb<sup>1</sup> of pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Journal of High Energy Physics **2019** no. 9, (2019) 91, [https://doi.org/10.1007/JHEP09\(2019\)091](https://doi.org/10.1007/JHEP09(2019)091).
- [108] ATLAS Collaboration, *Variable Radius, Exclusive- $k_T$ , and Center-of-Mass Subjet Reconstruction for Higgs( $\rightarrow b\bar{b}$ ) Tagging in ATLAS*, Tech. Rep. ATL-PHYS-PUB-2017-010, CERN, Geneva, Jun, 2017. <https://cds.cern.ch/record/2268678>.
- [109] <https://keras.io/layers/recurrent/> (cit. on p. 65).
- [110] A. Collaboration and I. Caprini, *Performance of jet substructure techniques for large- $R$  jets in proton-proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, Journal of High Energy Physics **2013** (2013) .
- [111] G. Cowan and E. Gross, *Discovery significance with statistical uncertainty in the background estimate.*, Tech. Rep. ATLAS Statistics Forum, 2008. [www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf](http://www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf).
- [112] ATLAS Collaboration, *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, Journal of High Energy Physics **2014** no. 10, (2014), [http://dx.doi.org/10.1007/jhep10\(2014\)141](http://dx.doi.org/10.1007/jhep10(2014)141).
- [113] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at  $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Physical Review Letters **117** no. 18, (2016), <http://dx.doi.org/10.1103/physrevlett.117.182002>.
- [114] ATLAS Collaboration, *Combined search for the Standard Model Higgs boson in pp collisions at  $\sqrt{s} = 7$ TeV with the ATLAS detector*, Physical Review D **86** no. 3, (2012), <http://dx.doi.org/10.1103/physrevd.86.032003>.
- [115] L. Moneta, K. Cranmer, G. Schott, and W. VERKERKE, *The RooStats project*, Proceedings of 13th International Workshop on

- Advanced Computing and Analysis Techniques in Physics Research  
— PoS(ACAT2010) (2011), <http://dx.doi.org/10.22323/1.093.0057>.
- [116] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003.
- [117] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** no. 2, (2011), <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0>.
- [118] A. L. Read, *Presentation of search results: the CLs technique*, Journal of Physics G: Nuclear and Particle Physics **28** no. 10, (2002) 2693–2704, <https://doi.org/10.1088>
- [119] ATLAS Collaboration, *Standard Model Summary Plots Summer 2019*, Tech. Rep. ATL-PHYS-PUB-2019-024, CERN, Geneva, Jul, 2019. <http://cds.cern.ch/record/2682186>.