Università degli Studi di Napoli Federico II

Dipartimento di Fisica "Ettore Pancini"

Dottorato di ricerca in Fisica
Ciclo XXXII
Coordinatore: prof. Salvatore Capozziello

Machine Learning methods for diboson searches in semi-leptonic final states with the ATLAS experiment at LHC

Settore Scientifico Disciplinare Fis/01

Dottorando:
Antonio Giannini

Tutors:
Prof. Leonardo Merola
Dott. Gianpaolo Carlino
Dott.ssa Elvira Rossi

Anni 2017/2020

To my family

Phileas Fogg aveva dunque vinto le ventimila sterline, ma, poiché durante il viaggio ne aveva speso circa diciannovemila, il risultato pecuniario era mediocre, anche se, come abbiamo detto, ***l'eccentrico gentleman non aveva cercato, in quella scommessa, nient'altro che la lotta, non la ricchezza****. Inoltre, le mille sterline che rimanevano le spartì tra l'onesto Gambalesta e il malaugurato Fix, cui non sapeva serbare rancore; solo che, per la regolarità, defalcò al suo servitore l ammontare delle millenovecentoventi ore di gas spese per colpa sua.*

...

Sicché Phileas Fogg aveva vinto la scommessa: aveva compiuto in ottanta giorni quel viaggio intorno al mondo. A tal fine aveva impiegato tutti i mezzi di trasporto: piroscafi, treni, carrozze, yachts, navi mercantili, slitte, elefanti. In tale occasione l'eccentrico gentleman aveva palesato le sue mirabili doti di sangue freddo e di precisione. ***E poi? Che cosa aveva guadagnato da quella serie di trasferimenti? Quale profitto aveva tratto da quel viaggio? Niente direte. Ebbene, niente, tranne una moglie deliziosa, che - per quanto ciò possa sembrare inverosimile - fece di lui il più felice degli uomini. Diciamo la verità: chi di noi, per un compenso anche inferiore, non farebbe il Giro del Mondo?***

Il giro del mondo in 80 giorni, cap XXXVII, 1873, Jules Verne

Contents

Introduction

My Thesis work has been devoted to a *Machine Learning* approach to analyze diboson events (ZZ, ZW or WW) in semi-leptonic decay channels. These final states could be used for searches of new resonances and for measurements of the cross-section in the non-resonant interpretation.

Nowdays, the high energy particles physics is well described inside a theoretical model, known as the Standard Model (SM) of particle physics [1], that describes three of the fundamental interactions, strong, electromagnetic and weak. The success of the SM in describing the wide range of precise experimental measurements is a remarkable achievement. The SM is constructed from a number of beautiful and profound theoretical ideas put together in order to reproduce the experimental data.

Several precision measurements, performed in the last decades of experiments, proved with a remarkable level of precision all the SM predictions. The most recent confirmation has been the discovery of a new particle with a mass of $125\ GeV$, announced on the 4^{th} *July* 2012. Following results probed that this new particle is compatible with the Higgs boson theorised as the last brick to complete the SM. Recently, one of the rarest SM process that can be produced and observed at the LHC, has been claimed; this is the *Vector Boson Scattering* that is strictly related to the presence of the Higgs boson and its role in the SM.

Despite this success there are many unanswered questions, motivating beyond SM (BSM) searches. It is a widely held opinion within the scientific community that the SM is an effective theory which we currently probe at low energy, while a general theory will begin to become accessible when the predictions of the SM start to become incorrect. Several theories beyond the SM exist, these theoretically well motivated extensions include the Heavy Vector Triplet model or the Randall-Sundrum models with warped extra dimensions. In particular, these theories are predicted to manifest themselves with the presence of heavy resonances decaying to vector boson pairs.

The diboson topology has been at the centre of this thesis work; the semi-leptonic decay channel has been investigated to due the clean experimental signature of a boson decaying in leptons and the sufficiently high decay rate of the other boson in hadrons. These decay channels permit searches of new resonances and also measurements of the anomalous coupling and cross-section in the non-resonant interpretation.

Depending on the assumed model, events with 2 extra jets in the final state coming from the Vector Boson Fusion production mechanism in the resonant case and from Vector Boson Scattering in the non-resonant one can be produced. The study of these topologies is a challenging topic for probing new Physics. My work during this Thesis has been the study of these channels and the increase of the sensitivity to observe them. My main contribution to these analyses has been the definition of a Machine Learning approach to improve the efficiency in the identification of the production mechanism of the dibosonic couple. Machine Learning algorithms and deep learning show improvement in the classification problems that are faced up at the experiments, motivating more efforts in the development of these techniques. In my Thesis, the new Recurrent Neural Network is used as powerful tool for the identification and the selection of the production mechanism of the new resonances predicted in BSM models, such as Vector Boson Fusion mechanism, while a BDT has been used for the Vector Boson Scattering in the non-resonant channel. The analysis described has the purpose to investigate a mass range coming from 300 GeV up to 5 TeV using data from proton-proton collisions at a center of mass energy of $\sqrt{s} = 13\ TeV$, corresponding to an integrated luminosity of 139 fb^{-1} recorded with the ATLAS detector from 2015 to 2018 at the Large Hadron Collider.

The Thesis consists of six chapters. Chapter 1 summarizes the key concepts and of the SM and provides a description of the BSM models and of the main background processes that are tested in this thesis. Chapter 2 presents the main characteristics of LHC and experiment. Chapter 3 gives a description of the main physics objects and of the reconstruction algorithms that bring the information of the detectors to physics analyses. Chapter 4 shows an overview of Machine Learning algorithms. Chapter 5 describes the analysis details focusing on the features of the Machine Learning methods used. Finally, in Chapter 6 the results of the resonant search using the full run-II dataset and of the cross section measurement of the non-resonant interpretation with a partial dataset of the run-II are showed.

Chapter 1

The Standard Model and beyond

A general overview of the theoretical framework of particle physics is provided in this chapter. The basic concepts of the Standard Model (SM) of elementary particles are given, from the gauge invariance theories to the spontaneous symmetry breaking. Then, the attention will focus on the main physics processes that occur at a hadron collider, moving, finally, to an overview of the theoretical models of Physics beyond the SM.

1.1 Introduction

The aim of the particle physics is to described the fundamental components of matter and their interactions. The SM of particle physics describes the electromagnetic, strong, and weak interactions and provides the current best explanation of phenomena we observe in the high energy physics.

The formulation of the SM is based on three main ideas [1]:

- The *parton model*, proposed in the 1964 independently by Gell-Mann and Zweig [2]-[3], in which hadrons are made of quarks and anti-quarks, combined within a non-dynamic model, that allowed to describe the compositions of known hadrons and to predict the existence of new ones. The idea was supported by the experiment carried out at SLAC in the 1968 [4]-[5]; the electron-proton scattering at large angles showed that the protons are made of point-like particles.
- The *gauge principle* allowed Yang and Mills to construct a gauge theory of strong interaction [6].
- The *spontaneously symmetry breaking* exploited the possibility that there might be symmetries of a physics system that are not symmetries of the vacuum state of the system. This is the key idea of the developing of the Higgs mechanism [7, 8] that predicts a new boson particle, called the Higgs boson, finally observed in the 2012.

1.2 Gauge principle

The four fundamental interactions can be described in terms of gauge theories. Indeed, all are derived from a principle, the *gauge principle*, introduced by H. Weyl in 1929 [9]. Often physics theories describing dynamical systems can be derived using *symmetry principles*. A symmetry is a mathematical transformation that, when applied to the equations that describes a physics system, does not change the physics observables of the system.

When a Lagrangian is invariant under a symmetry transformation it is *gauge invariant*. The "gauge" term indicates that the Lagrangian contains more degrees of freedom with respect to the physics ones. A Lagrangian can have global and/or local gauge invariance. Global invariance is referred to an invariance that does not depend on a particular space-time coordinate. A local invariance of the transformation acts in different ways on different space-time coordinate. In general, if the Lagrangian that describes a physics system is globally invariant is not necessarily also locally invariant. Gauge invariance is discussed in the context of classical quantum mechanics in the next Section.

1.2.1 Gauge invariance in quantum mechanics

The Schrödinger equation for a particle with charge q and mass m moving in an electromagnetic field is:

$$\left(\frac{1}{2m}(-i\nabla - q\vec{A})^2 + qV\right)\psi(\vec{x},t) = i\frac{\partial\psi(\vec{x},t)}{\partial t}. \tag{1.1}$$

The general solution $\psi(\vec{x},t)$ of the equation completely describes the dynamics of the particle. As general result of classical electro-magnetics fields, these potentials are not unique but they can change under a gauge transformation:

$$\begin{aligned} V \to V' &= V - \frac{\partial\chi}{\partial t} \\ \vec{A} \to \vec{A}' &= \vec{A} + \nabla\chi. \end{aligned} \tag{1.2}$$

The resulting Maxwell equations for the physics fields $\vec{E}$ and $\vec{B}$ will not change. Applying a gauge transformation to the electro-magnetics potentials, the new equation will have a solution $\psi'(\vec{x},t) \neq \psi(\vec{x},t)$; this means that the equation 1.1 is not gauge invariant unless a change in ψ, $\psi'(\vec{x},t) \to \psi(\vec{x},t)$, is allowed when a gauge transformation is applied to the Maxwell potentials. This is possible since the function ψ is not a direct observable but only the $|\psi|^2$ is a real number interpreted as the probability density of measuring a particle as being at a given place at a given time or having a definite momentum. In this way the form of the equation after the transformation:

$$\left(\frac{1}{2m}(-i\nabla - q\vec{A}')^2 + qV'\right)\psi'(\vec{x},t) = i\frac{\partial\psi'(\vec{x},t)}{\partial t} \tag{1.3}$$

is exactly the same, that means the "same physics" is described before and after the transformation. The form of the transformation for ψ can be derived from the gauge transformation of the potentials:

$$\psi'(\vec{x},t) = e^{iq\chi(\vec{x},t)}\psi(\vec{x},t) \tag{1.4}$$

where χ is the same space-time dependent function that appears in the gauge transformation of the potentials **??**. This implies the equations:

$$\begin{aligned} -i\vec{D}'\psi' &= e^{iq\chi}(-i\vec{D})\psi \\ iD'^0\psi' &= e^{iq\chi}iD^0\psi \end{aligned} \tag{1.5}$$

where the *covariant derivative* has been introduced:

$$\begin{aligned} \vec{D} &= \nabla - iq\vec{A} \\ D^0 &= \frac{\partial}{\partial t} + iqV \end{aligned} \tag{1.6}$$

With this transformations the gauge invariance of the Maxwell equations re-emerge as covariance in quantum mechanics provided that the combined transformations of the potentials and of the wave function are made. The Schrödinger equation is non-relativistic, but the prescriptions and the conclusions showed here are applicable in the same way at the equations of quantum electrodynamics.

1.2.2 Gauge principle for quantum electrodynamics (QED)

In the previous section, the Schrödinger equation for a charged particle in an electromagnetic field was written and its gauge invariance under the combined transformations of the electromagnetic potentials and of the wave function were shown. In this section, the argument will be reversed: requiring that the theory is invariant under the space-time-dependent phase transformation (gauge principle) it is possible to build the interaction theory of the electromagnetism starting from a free particle theory.

The theory of free Fermi fields is described by the following Lagrangian density [1] :

$$\mathcal{L} = \bar{\psi}(x)(i\partial\!\!\!/ - m)\psi(x) \tag{1.7}$$

it can be demonstrated that is invariant under a global $U(1)$ transformation $\psi(x) \longrightarrow e^{i\alpha}\psi(x)$ [2] , where α is a constant parameter; this is true because the derivative term transforms as:

$$\partial_\mu\psi(x) \longrightarrow e^{i\alpha}\partial_\mu\psi(x) \tag{1.8}$$

The same argument can not be applied if the $U(1)$ transformation is required to be local, $\alpha = \alpha(x)$:

$$\psi(x) \longrightarrow e^{i\alpha(x)}\psi(x) \tag{1.9}$$

therefore:

$$\partial_\mu\psi(x) \longrightarrow e^{i\alpha(x)}\partial_\mu\psi(x) + ie^{i\alpha(x)}\partial_\mu\alpha(x)\psi(x) \tag{1.10}$$

the gauge principle requires the local gauge invariance of the theory; this is possible introducing a covariant derivative D_μ instead of the ∂_μ such that:

$$D_\mu\psi(x) \longrightarrow e^{i\alpha(x)}D_\mu\psi(x) \tag{1.11}$$

this represent the only solution in order to guarantee the invariance of the theory under the chosen group.

The form of the covariant derivative can be written in the generic form depending by the usual derivative operator an a field $A^\mu(x)$:

$$D_\mu = \partial_\mu + igA_\mu(x). \tag{1.12}$$

The Lagrangian is invariant under the local transformation, so the gauge principle is satisfied, if the introduced field $A(x)$ transforms as:

$$A_\mu(x) \longrightarrow A_\mu(x) - \frac{1}{g}\partial_\mu\alpha(x) \tag{1.13}$$

In summary, requiring an invariance of the theory under a local $U(1)$ symmetry has:

- promoted a free theory of fermions to an interacting theory between the particle field ψ and the force field A_μ
- fixed the form of the interaction in terms of a new vector field A_μ:

$$\mathcal{L}_{int} = -g\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x) \tag{1.14}$$

- no mass term $A_\mu A^\mu$ is allowed by the symmetry.

[1] In the following the x represents the space-time point in the Minkowski space with Lorentz metrics, $\mathbb{R}^4_1$

[2] The unitary group of degree n, denoted U(n), is the group of $n \times n$ unitary matrices, with the group operation of matrix multiplication.

The last point is not a limitation of the theory if this concept is used to build the *Quantum Elecro Dynamics* (QED) since the mediator of that force is massless.

Furthermore, a kinetic term could be added to the Lagrangian:

$$-\frac{1}{4}F^{\mu\nu}(x)F_{\mu\nu} \qquad \textit{with} \qquad F_{\mu\nu} = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \tag{1.15}$$

that is invariant under the local transformation of $U(1)$.

1.3 Standard Model of particle physics

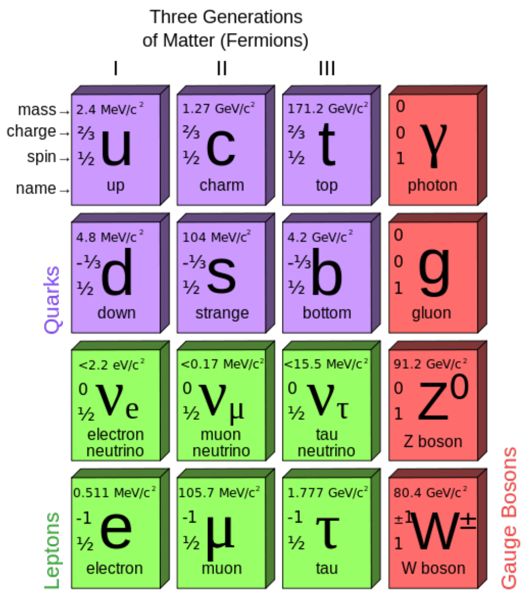
The fundamental interactions observed in nature are four: strong, electromagnetic, weak and gravitational.

The reduced strength of the gravitational interaction produces negligible effects on the behaviour of sub-atomic particles, and plays no role in determining the internal properties of matter.

The Quantum Field Theory represents the fundamental formal and conceptual framework of the SM, a gauge theory that describes both particles and forces in term of quantum fields. The fundamental particles are organised in two categories according to the spin, as shown in Table 1.1:

- the fermions are particles that obey to the Fermi-Dirac statistics; they are half-integer spin particles and they are the elementary pieces of ordinary matter;
- the bosons are particles that obey to the Bose-Einstein statistics; they are integer spin particles and they are the particles that describes the fundamental forces according the quantisation of the fields of the forces.

Figure 1.1 show a picture of the puzzle of the fundamental particles, known nowdays.



(a)

Figure 1.1: Puzzle of the Standard Model particles.

The Standard Model is the model that describes the behaviour of the elementary particles and of the weak, electromagnetic and strong interactions they are sensitive to; the SM is build starting from the gauge principle, requiring the theory to be invariant for a local gauge symmetry group:

$$SU(3)_{col} \otimes SU(2)_L \otimes U(1)_Y. \tag{1.16}$$

	Generation			Quantum numbers			
	1	2	3	I	I_3	Y	Q[e]
Leptons	$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}_L$	1/2 1/2	1/2 −1/2	−1 −1	0 −1
	e^-_R	μ^-_R	τ^-_R	0	0	-2	-1
Quarks	$\begin{pmatrix}u\\ d\end{pmatrix}_L$	$\begin{pmatrix}c\\ s\end{pmatrix}_L$	$\begin{pmatrix}t\\ b\end{pmatrix}_L$	1/2 1/2	1/2 −1/2	1/3 1/3	2/3 −1/3
	u_R	c_R	t_R	0	0	4/3	2/3
	d_R	s_R	b_R	0	0	-2/3	1/3

Table 1.1: Overview on the quantum numbers of the Standard Model fermions in the GWS model. The right-handed neutrinos don't take part to the SM interaction and they are not considered here.

This set of local transformations is a set of gauge transformations, meaning that to preserve symmetry under a given transformation it requires the introduction of additional fields via the application of the gauge principle (1.2.2). These are spin 1 gauge fields in the Lagrangian, that are associated with new particles. In particular, $SU(3)_{col}$ is a non abelian gauge symmetry group which describes the strong interactions. The generators of the symmetry group are eight. This geometrical relation reflects the fact that the strong interaction is carried by eight massless particles, the *gluons*. In this interpretation the gluons are the mediators or the strong interaction and they have a strong charge, known as "color". Gluons and quarks strong interactions are described by the Quantum Cromodynamic (QCD) theory. $SU(2)_L \otimes U(1)_Y$ is the weak isospin symmetry group, introduced by Glashow-Weinberg-Salam (GWS) [10]-[11]-[12], which describes the unified electromagnetic and weak (EWK) interactions. The mediators of the EWK interaction are three massive vector bosons W^+, W^- and Z^0, plus the photon, γ.

The elementary particles are organised in multiplets of the fundamental groups that represents the symmetry of the specific interaction. The leptons, interacting only via the electromagnetic (γ boson) and weak forces ($W^\pm/Z^0$ bosons), are doublets of the $SU(2)_L$ symmetry group, while the quarks, interacting also via the strong force (gluons), are weak doublets with both components being triplets of the $SU(3)_c$ group:

$$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L, \qquad \begin{pmatrix}u\\ d\end{pmatrix}_L \qquad \Longleftrightarrow \qquad \begin{pmatrix}\color{red}{u} & \color{green}{u} & \color{blue}{u}\\ \color{red}{d} & \color{green}{d} & \color{blue}{d}\end{pmatrix}_L \tag{1.17}$$

1.3.1 Higgs sector

The Lagrangian of the unbroken $SU(2)_L \otimes U(1)_Y$ gauge theory of vector bosons and fermions describes all the partilces of the SM as massless. The mediating particles (except for the photon) neither the fermions in the electroweak sector are massless according to the experiments. It is possible to introduce the mass terms in the theory via a mechanism that does not break the gauge symmetry of the Lagrangian.

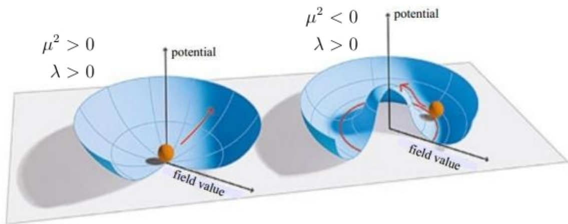
The mechanism of *spontaneous electroweak symmetry breaking* applied to a *non-abelian* theory was introduced by Peter Higgs [7] in 1964, and independently by Robert Brout and Francoise Englert [8], and Gerald Guralnik, C. R. Hagen, and Tom Kibble[13],[14]. It provides a solution to the massless fields and it is commonly known as the *Higgs mechanism*.

According to this mechanism, some additional fields are introduced with a potential which causes the vacuum (the ground state) to break the symmetry spontaneously.

In the simple case of an abelian gauge theory with a vector field $A_\mu(x)$ coupled to one complex scalar field $\phi(x)$ it is possible to introduce an interaction term through:

$$\mathcal{L}_\phi = (D^\mu\phi)^* D_\mu\phi - V(\phi) = (D^\mu\phi)^* D_\mu\phi - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2 \tag{1.18}$$

where the $V(\phi)$ is the know $\lambda\phi^4$ scalar theory and covariant derivative $D_\mu = \partial_\mu + igA^\mu$. This lagrangian is invariant under the $U(1)$ symmetry. Any mass term for the field A_μ would break the $U(1)$ gauge invariance. The introduced potential $V(\phi)$ has functional shapes showed in Figure 1.2.



(a)

Figure 1.2: Representation of the shape of the Higgs potential $V(\phi)$ in the case $\mu^2 > 0$ (a) and $\mu^2 < 0$ (b).

Two different cases are possible according to the sign of the parameter μ^2:

- $\mu^2>0$: the theory describes the case of the electrodynamics of a mass-less photon ($g = -e$) and a massive scalar field of mass μ;
- $\mu^2<0$: we have a theory in which both the scalar and the vector fields are massive once a particular minima is chosen; in this case, the original $U(1)$ symmetry of the lagrangian is said to be spontaneously broken or hidden:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}g^2v^2A_\mu A^\mu + \frac{1}{2}(\partial^\mu\phi_1)^2 + \mu^2\phi_1^2 + \frac{1}{2}(\partial^\mu\phi_2)^2 + gvA_\mu\partial^\mu\phi_2 \tag{1.19}$$

where v is related to the value of the field at the minimum through $\phi_0 = (-\mu^2/2\lambda)^{1/2} = v/\sqrt{2}$ and the $\phi_{1/2}$ are the real and complex component of the field after the choice of the minimum:

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}}\big(\phi_1(x) + i\phi_2(x)\big) \tag{1.20}$$

Usually, a more convenient parametrisation (*unitary gauge*) is used:

$$\phi(x) = \frac{e^{i\frac{\chi(x)}{v}}}{\sqrt{2}}(v + H(x)) \qquad \xrightarrow[U(1)]{} \qquad \frac{1}{2}(v + H(x)) \tag{1.21}$$

the χ degree of freedom represent a Goldstone boson and it "disappears" under the $U(1)$ invariance and the lagrangian become:

$$\mathcal{L} = \mathcal{L}_A + \frac{g^2v^2}{2}A^\mu A_\mu + \frac{1}{2}\big(\partial^\mu H\partial_\mu H + 2\mu^2H^2\big) + ... \tag{1.22}$$

that now describes the dynamics of a system made of:

- a massive vector field A^μ with $m_A^2 = g^2v^2$;
- a real scalar field H of mass $m_H^2 = -2\mu^2 = 2\lambda v^2$, the Higgs field.

during the spontaneous symmetry breaking procedure the number of degree of freedom has been preserved. Indeed, the theory had four degree of freedom, given by the two coming from the vector mass-less field and the two coming from the massive complex scalar field, after the symmetry breaking there are still four degrees of freedom, three from the massive vector field and one from the scalar and real field.

In the specific case of the SM, in which there are several vector fields $A^a_\mu(x)$, it is convenient to use several real and scalar fields $\phi_i(x)$. This time the interactive term between the scalar field and the vector fields, that will generate the mass terms after the spontaneous symmetry breaking, are:

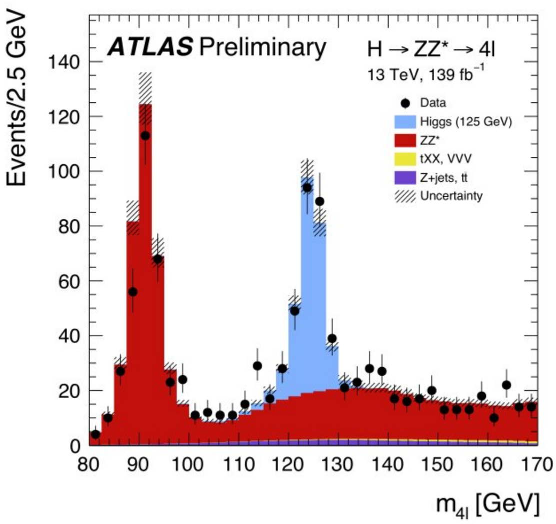
$$\frac{1}{2}(D_\mu\phi)^2 \longrightarrow ... + \frac{1}{2}g^2(T^a\phi)_i(T^b\phi)_i A^a_\mu A^{b,\mu} + ... \tag{1.23}$$

$$\xrightarrow[\phi_{min}=\phi_0]{} ... + \frac{1}{2}g^2(T^a\phi_0)_i(T^b\phi_0)_i A^a_\mu A^{b,\mu} + ... \tag{1.24}$$

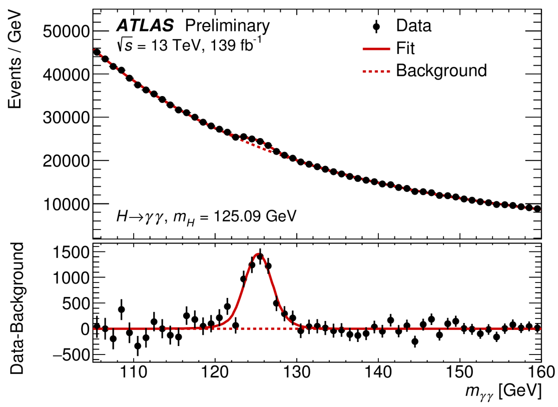
the matrix $g^2(T^a\phi_0)_i(T^b\phi_0)_i$ represents the mass terms introduced after the breaking of the symmetry; in particular:

- $T^a\phi_0 \neq 0$: massive vector boson + Goldstone boson;
- $T^a\phi_0 = 0$: massless vector boson + massive scalar field.

A new particle with mass $\sim 125\ GeV$ has been observed during the Run-I data taking of the LHC from both ATLAS and CMS collaboration with the announcement given on the *July* 4^{th} 2012. All the studies done show the compatibility of that new particle with the SM Higgs booson. The "discoveries" plots with the full Run-II statistics are showed in Figures 1.3.



(a)



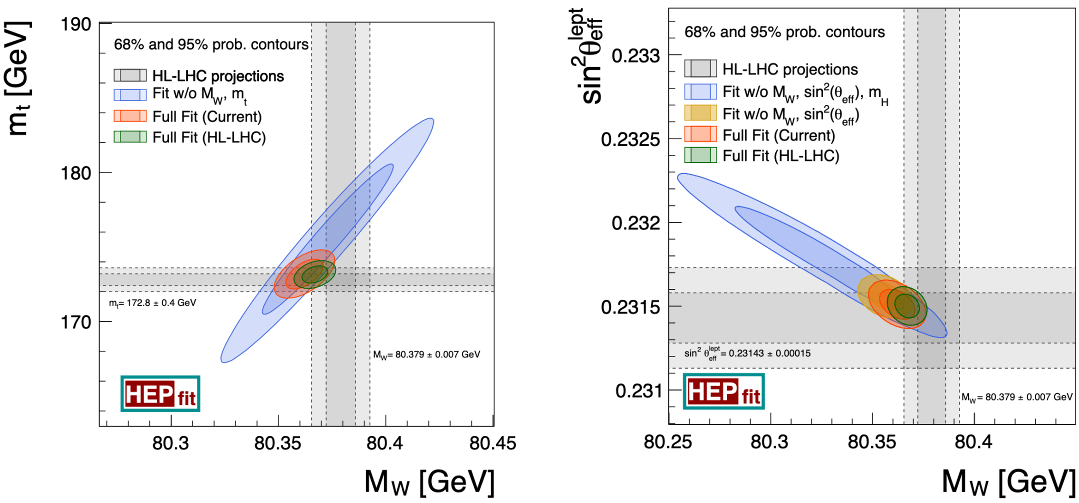
(b)

Figure 1.3: Invariant mass distribution of the Higgs boson observed in the $H \to 4l$ and $H \to \gamma\gamma$ decay channel with the full Run-II dataset.

1.3.2 Electroweak Symmetry Breaking phenomena at LHC

In the SM physics the aim of the theoretical and experimental searches is to probe the *ElectroWeak Symmetry Breaking* (EWSB) with precision physics measurement at the energy frontier of the LHC. Precision is intrinsic to having a predictive theory like the Standard Model could be. Precision is effective when both theory and experiments have a way to reach comparable accuracy and improve it systematically. Particle physics has a very successful history of constraining new physics through precision measurements like precision fits of electro-weak observables (LEP, SLD, Tevatron, LHC), and the indirectly constrained on the Higgs mass [15]. Before the Higgs boson discovery, the EW fit already gave a very good prediction on the Higgs boson mass using the measurement on the all the other EW observable; the constraint on the Higgs boson mass was $m_H < 152\ GeV$.

The fit of the EW precision observable (EWPO) represent a fully stress-testing of the SM consistency [15]. The current results 1.4 show a very good agreement between indirect determination of EWPO and experimental measurements, also perspective to the HighLumi-LHC program is computed; this represents a strong constraint on any physics beyond the SM.



(a) (b)

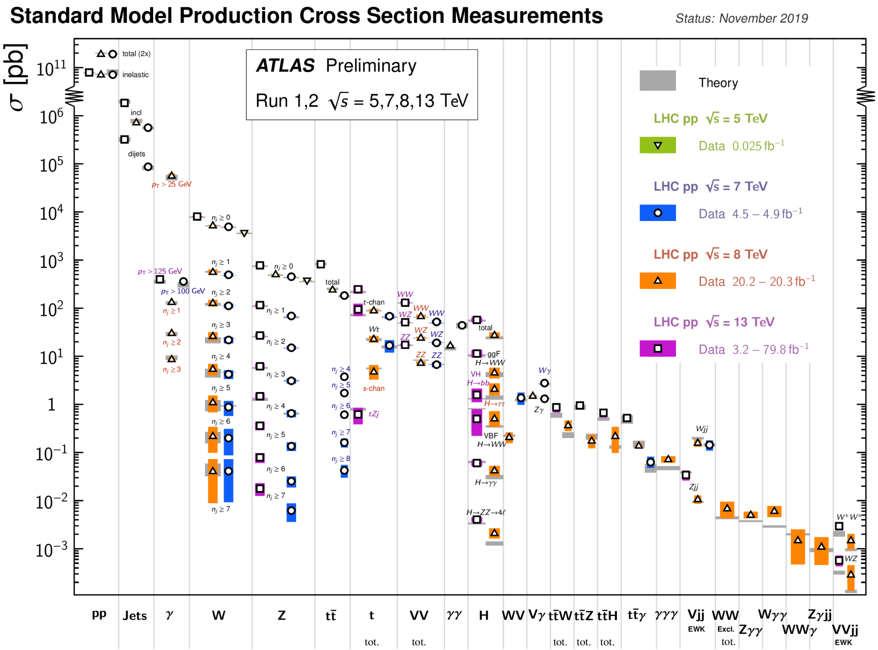
Figure 1.4: EWPO fit in different scenario; in particular the HL-LHC projection is also computed.

From Run 1 to Run 2 and beyond, a crucial point has been to develop the LHC Higgs precision program, in terms of precision measurement of Higgs properties (mass, couplings, width) and in terms of constraints on anomalous interactions. A general aim is to explore indirect evidence of new physics while searching for direct one and perform precision measurements. This has been one the aim of this thesis performing the cross section measurement of diboson process production via the *Vector Boson Scattering (VBS)*, that will be introduced in Section 1.4.4; this process arises from the spontaneous symmetry breaking mechanism described in Section 1.3.1.

The measurement of the Higgs boson mass in Run1 and improved in Run2 [16] promoted the SM parameter M_H to be an EW precision observable to be used in the global EW fit. Effects of New Physics can now be more clearly disentangled in both EW observables and Higgs-boson couplings, probing EWSB. Moreover, from decays $H \to VV$ and $H \to f\bar{f}$ it is possible to constrain more the spin and the parity of the Higgs boson.

1.4 SM processes at hadron colliders

When a proton to proton collision occurs several processes are allowed within the SM. The quarks interacts via the electromagnetic, strong and weak forces, foreseen final states with a large set of particles and production cross sections that vary on a very large range, as showed in Figure 1.5. Following the main processes that occur at pp collisions and that have been subject of this thesis will be introduced.



(a)

Figure 1.5: Summary of several Standard Model total and fiducial production cross section measurements, corrected for branching fractions, compared to the corresponding theoretical expectations.

1.4.1 V+ jets

The production of a Z/W boson in association with jets is the most SM processes that involves EW boson that occurs at the pp colliders [17]. The measurement of the production of this process constitutes a powerful test of perturbative quantum chromodynamics (QCD) [18, 19]. The large production cross section and easily identifiable decays of the Z boson to charged leptonic final states offer clean experimental signatures which can be precisely measured. Such processes also constitute a non-negligible background for studies of the Higgs boson and in searches for new phenomena; typically in these studies, the multiplicity and kinematics of the jets are exploited to achieve a separation of the signal of interest from the SM $V + jets$ process. These quantities are often measured in control regions and subsequently extrapolated to the signal region with the use of Monte Carlo (MC) generators, which are themselves subject to systematic uncertainty and must be tuned and validated using data. Predictions from the most recent generators combine next-to-leading-order (NLO) multi-leg matrix elements with a parton shower(PS) and a hadronisation model. Fixed-order parton-level predictions for V+jets production at next-to-next-to-leading order (NNLO) are also available [20, 21].

1.4.2 SM Diboson

The SM d-boson production at the pp colliders foresees the combination of the two EW gauge bosons, $W^{\pm}, Z^0$, so, three different flavours di-boson states are possible, ZZ, WZ and WW. The study of $W^{\pm}/Z^0$ diboson production is an important test of the Standard Model (SM) for its sensitivity to gauge boson self-interactions, related to the non-Abelian structure of the electroweak interaction [22]; this channel is also used to measure the boson polarization.

In pp collisions at a center-of-mass energy of $\sqrt{s} = 13\ TeV$, ZZ production [23] is dominated by quark–antiquark ($q\bar{q}$) interactions, with an O(10%) contribution from loop-induced gluon–gluon (gg) interactions [24, 25]. The SM ZZ production can proceed via a Higgs boson propagator, although this contribution is suppressed in the region where both Z bosons are produced on-shell. As such, non-Higgs ZZ production is an important background in studies of the Higgs boson [26]. It is also a background in searches for new physics producing pairs of Z bosons at high invariant mass [27] and sensitive to triple neutral-gauge-boson couplings, which are not allowed in the SM [28].

1.4.3 Top quark

The top quark is the heaviest elementary particle in the Standard Model (SM), with a mass m_t of around 172.5 GeV [29], which is close to the electroweak symmetry breaking scale. At the LHC, top quarks are primarily produced in quark-antiquark pairs ($t\bar{t}$), and the precise prediction of the corresponding inclusive cross-section presents a substantial test for perturbative QCD calculation techniques. The $t\bar{t}$ production rate may also be modified due to Beyond the Standard Model (BSM) physics. Theoretical predictions of the $t\bar{t}$ cross-section are available at next-to-next-to leading order (NNLO) in QCD including the resummation of the next-to-next-to-leading logarithmic (NNLL) soft-gluon terms [30, 31]. Measurements of $\sigma_{t\bar{t}}$ at 7, 8 and 13 TeV in pp collisions have been performed by both ATLAS [32, 33, 34] and CMS [35, 36] collaborations. Additionally, the CMS Collaboration measured the $t\bar{t}$ cross-section at $\sqrt{s} = 5.02\ TeV$ [37].

1.4.4 Vector Boson Scattering

After the discovery of the Higgs boson, the scrutiny of the electroweak symmetry breaking (EWSB) becomes a main focus at the LHC. In addition to direct measurements of Higgs boson properties, the study of massive vector-boson scattering (VBS) offers another key avenue to probe the EWSB [38, 39, 40]. In the Standard Model (SM), the Higgs boson acts rigorously to prevent longitudinal VBS amplitudes from violating unitarity at the TeV scale [38]; therefore, the study of high-energy behaviours of VBS is crucial to understand the nature of the EWSB. Many new physics scenarios, such as Supersymmetry and Little Higgs models [41], offer alternative EWSB mechanisms, which can manifest themselves as appearance of heavy particles or modifications of Higgs couplings in the accessible energy regime. These new phenomena could manifest themselves in rises of amplitudes or resonant structures in the TeV range and thereby alter the way of delicate cancellations at high energies.

While no VBS processes were established prior to the LHC era, the LHC provides an unprecedented opportunity to study them, owing to its high collision energies and luminosities. At the LHC, VBS is typically produced as two vector bosons radiated from initial-state quarks and then scattering into another pair of vector bosons. The detector signature of VBS contains decay products of the pair of outgoing bosons and a pair of hadronic jets, hereafter denoted as $VVjj$. The most promising channel to measure VBS is therefore the electroweak (EW) production of $VVjj$ ($EWVVjj$), which has no quantum chromodynamics (QCD) vertices at leading order (LO). The QCD production of $VVjj$ contains two QCD vertices at the lowest order (denoted as $QCDVVjj$ processes) and constitutes an irreducible background to the searches for $EWVVjj$ production. The characteristics of $EWVVjj$ production include a large separation in rapidity between the two jets ($\Delta y(jj)$) as well as a significant invariant mass of the jet pair (m_{jj}), which are often utilized to improve the signal-to-background ratio.

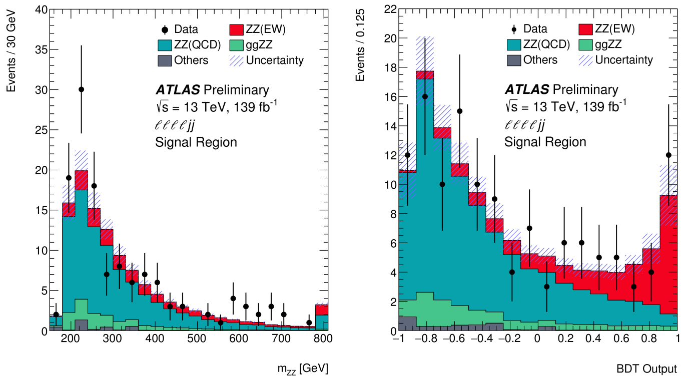


Figure 1.6: Observed and expected m_{ZZ} (a) and multivariate discriminant distributions after the statistical fit in the CRs. The error bands include the experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty on data.

The search for electroweak production of two Z boson in association with two jets in the full leptonic final state, $llll, \nu\nu ll$, with the full run-II dataset of the ATLAS collaboration gives the first observation of this process [42]. The reached significance in the background-only hypothesis has been of 5.5σ. The measured cross-section for electroweak production in the fiducial region is $0.82 \pm 0.21\ fb$, corresponding to a signal strength of 1.35 ± 0.34, in agreement with the SM prediction.

1.5 Beyond the Standard Model

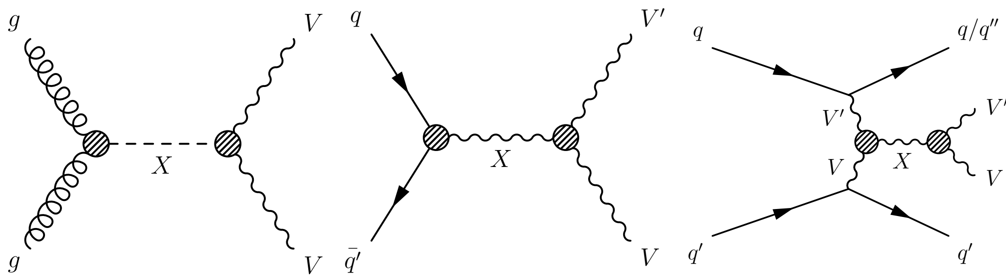
The SM suffers from several theoretical problems, despite being able to explain with high precision most of the experimental data that has been produced until now:

- no dark matter candidate is provided by the SM;
- does not explain the gravitational interaction;
- the level of CP violation is not sufficient to explain the matter anti-matter asymmetry seen in the universe;
- it does not explain the hierarchy problem, i.e. why gravity is so weak compared to the other interactions;
- fine tuning is required to deal with divergences in the Higgs sector.

Because of these reasons and indeed others not discussed, it is a widely held opinion within the scientific community that the SM is an effective theory which we currently probe at low energy. The general theory will begin to become accessible when the predictions of the SM start to become incorrect. More precise determination of the free parameters of the SM will allow the scale at which this happens to be better understood. Several beyond the SM theories exist which describe the SM predictions at low energy.

1.5.1 New heavy particle decaying into diboson final states

Many models beyond the Standard Model of particle physics predict heavy particles that could decay into diboson final states. Production through gluon–gluon fusion (ggF), Drell–Yan (DY) and vector-boson fusion (VBF) processes are considered, depending on the assumed model. Representative Feynman diagrams of these processes are shown in Figure 1.7. Below I will describe a subset of the models predicting a heavy diboson resonance.



(a) *gluon-gluon fusion* (b) *Drell-Yan* (c) *Vector Boson Fusion*

Figure 1.7: Representative Feynman diagrams for the production of heavy resonances X with their decays into a pair of vector boson.

1.5.2 Heavy Vector Triplet model

In the *Heavy Vector Triplet model* (HVT), couplings of new heavy particles to quarks, leptons, SM gauge and Higgs bosons are present, further the new triplet has zero hypercharge.

The phenomenological features of such a triplet can be concisely described within the HVT setup introduced in Ref. [43]. An HVT consists of two essentially degenerate states: an electrically charged, $V^{\pm}$ (also denoted as W'), and a neutral one, V^0 (Z'). The couplings of the HVT to all SM particles are given in terms of the new coupling g_V, which parameterises the interaction strength between the heavy vectors. This makes the setup extremely versatile since it can capture the features of many, weakly and strongly coupled, concrete models. For illustration, two explicit models are chosen as benchmarks, First, an extended gauge symmetry discussed in Ref. [44] is used as a benchmark of a weakly coupled model, referred to as model A in the following. The value of the coupling g_V is small, of order one, in these models. Second, we consider the strongly interacting Minimal Composite Higgs Model [45], referred to as model B in the following, which is valid for larger values of g_V. Values within the range $1 \lesssim g_V \lesssim 4\pi$ are acceptable. However, since the width of the HVT grows with g_V in model B, large values of g_V do not produce a narrow resonance. In this search we focus on narrow resonances and therefore consider values of g_V in the range $1 \lesssim g_V \lesssim 6$ for which Γ/M never exceeds around 10%.

The relevant parameter space of an HVT with a given mass is two-dimensional consisting of two parameter combinations which describe its couplings to fermions and to SM gauge bosons. The HVT model considered in this work is a simplified version with an universal coupling C_F of V to fermions. The coupling of the HVT to fermions scales as $g_F = g^2 c_F / g_V$, where g is the SM $SU(2)_L$ gauge coupling and c_F is a free parameter which can be fixed in each explicit model. In both benchmark models A and B, c_F is expected to be of order one. The HVT coupling to fermions in strongly interacting models is thus g^2/g_V suppressed with respect to weakly coupled models. Thus, in general, a large coupling g_V corresponds to a small Drell-Yan production rate and, similarly, a small branching ratio into fermionic final states. Concerning the HVT coupling to SM bosons, note that it couples dominantly to the longitudinal components of the gauge bosons and to the Higgs, while the coupling to transverse gauge bosons is generally suppressed. Contrary to the coupling to fermions, the HVT coupling to SM bosons scales as parameter the g_H, with $g_H = g_V c_H$. The parameter c_H, analogously to c_F, has to be fixed in each individual model and takes values of order one in models A and B. Here we have the reversed situation, that a small value of g_V in weakly coupled extensions of the SM leads to a small branching fraction into gauge bosons, while strongly coupled theories predict an enhanced branching ratio. 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 produces less, decay predominantly into gauge bosons and two-fermion final states can be extremely rare. The results can be presented as contours in the parameter space (g_H, g_F) which allow for a broad interpretation, going beyond the benchmark models A and B. In a very large class of explicit models of heavy vectors, the parameters c_H and c_F can be computed and the result compared with the aforementioned contours to asses the compatibility of the concrete model with the experimental search.

The two benchmark models, A and B, are predominately produced via quark-anti-quark annihilation. 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 $g_F = 0$.

Diboson final states, both neutral W^+W^-, ZH, and charged $W^\pm Z, W^\pm H$, where H is the SM Higgs boson, are particularly interesting in strongly coupled models where the branching ratio into diboson final states is enhanced. Note that the HVT coupling to two SM bosons comes from a gauge invariant coupling to the electroweak triplet Higgs current, with strength g_H, and thus all the couplings to the aforementioned final states are expected to be equal. In particular the HVT framework predicts the same branching ratios for the four processes:

$$\mathrm{BR}(V^\pm \to W^\pm Z) = \mathrm{BR}(V^\pm \to W^\pm H) = \mathrm{BR}(V^0 \to W^\pm W^\pm) = \mathrm{BR}(V^0 \to ZH)\,. \tag{1.25}$$

Other neutral diboson final states are either suppressed or forbidden. The decay of a spin-1 vector into HH is forbidden by momentum and angular momentum conservation (and consequently Lorentz invariance). This is true to all orders, so no higher dimensional operator can appear. ZZ is accidentally not present at dimension four. While operators can appear at dimension six and higher, they are highly suppressed and therefore not considered.

This relation is of primary importance in the HVT framework since it allows us to gain a higher sensitivity by combining not only neutral and charged channels, but also eventually channels involving the Higgs boson ([43]).

Note that the branching ratios of W' and Z' into bosons are the same. The reason is that the HVT W'/Z' both couple to the Higgs current from which the widths and branching ratios into SM gauge bosons are derived. These couplings are parametrically the same. Hence the widths into WW, WH and ZH are the same up to finite mass effects. These finite mass effects go as $m(W)/m(V)$ and are thus small in the relevant parameter space.

The HVT also couples to fermions. Here it is important to note that the triplet couples to the fermionic current. Therefore, what needs to be compared is the sum of the widths (or equivalently the sum of the BRs) of all quark and lepton final states. For example, the sum of the widths into $\ell\ell$ and $\nu\nu$ is the same as the width into $\ell\nu$. For the quark sector, the mixing has to be taken into account and the sum of all charged quark final states is equal to the sum of all neutral quark final states. Hence equation 1.25 is true in general.

1.5.3 Randal-Sundrum model

The Randall-Sundrum (RS1) framework [46] attempts to explain the hierarchy problem by introducing large extra dimensions in which SM fields can propagate. This leads to a tower of Kaluza-Klein (KK) excitations of SM fields, notably including KK excitations of the gravitational field that appear as $TeV - scale$ spin-2 Gravitons (G_{KK}) [47].

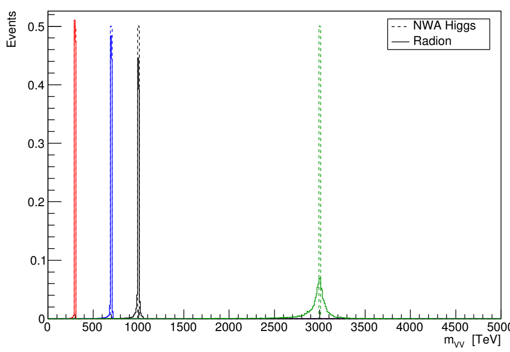
In some RS1 models the graviton has sizable couplings to all SM fields, which do not propagate significantly into the extra dimension (bulk). This leads to large production rates in both gluon-gluon (gg) and quark-quark (qq) fusion modes, and substantial decay rates to diphotons and dileptons. In the "bulk RS" scenario considered here, however, the SM fields are permitted to propagate into the bulk, where they are localised. The bulk RS model avoids the constraints on other RS scenarios arising from flavour physics and electroweak precision tests, at the cost of suppressing the couplings of the G_{KK} to light fermions, which leads to significantly reduced production rates from qq fusion and lower branching fractions to leptons and photons. The gg fusion production mode therefore dominates in the bulk RS model, with the G_{KK}-gluon coupling suppressed by a factor $\overline{M}_{\mathrm{Planck}}$, where k is the curvature scale of the extra dimension and $\overline{M}_{\mathrm{Planck}}$

is the reduced Planck mass. The value of $\overline{M}_{\mathrm{Planck}}$ is typically of order 1, and along with the mass of the G_{KK} is the only free parameter in this simplified model. The decays of the G_{KK} in this scenario are dominated by $G_{KK} \to t\bar{t}$, $G_{KK} \to HH$, and $G_{KK} \to V_L V_L$ (V_L is the longitudinally-polarized W and Z boson), with branching fractions that depend on mass.

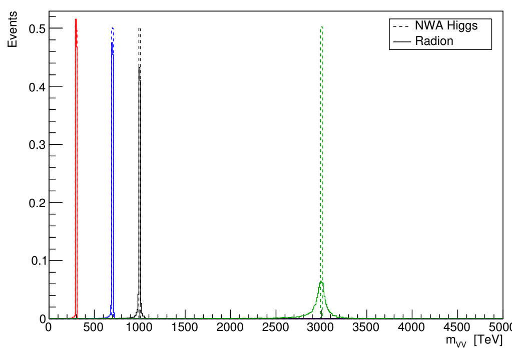
1.5.4 Radion model and Heavy Higgs

In the RS1 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 RS1 framework described above is that it lacks a mechanism to stabilise the radius of the compactified 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 [48, 49]. This cause the radion field to acquire a mass term, which is typically much smaller than the first KK excitation mass.

The coupling of the radion field to SM fields scale inversely proportional to the model parameter $\Lambda_R = \sqrt(g) \times k * e^{-k\pi r_c} \sqrt{\frac{M_5^3}{k^3}}$ where M_5 is the 5-dimensional Plank mass, which has been extensively studies in the literature[50, 51, 52]. 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. This lead to the dominant decay mode for the radion to be to pairs of bosons when the radion mass above $\sim 1\ TeV$. Both the production cross-sections for the radions and the total width scale like $\sim 1/\Lambda_R^2$. Phenomenologically, the radion has very similar couplings and production cross-sections as a heavy Higgs, but with much smaller widths. As shown in Figures 1.8-1.9-1.10-1.11, Radion signals have the same angular distributions as NWA Higgs and the mass width is smaller than the detector resolution (Breit-Wigner width is about 3% at 3 TeV).

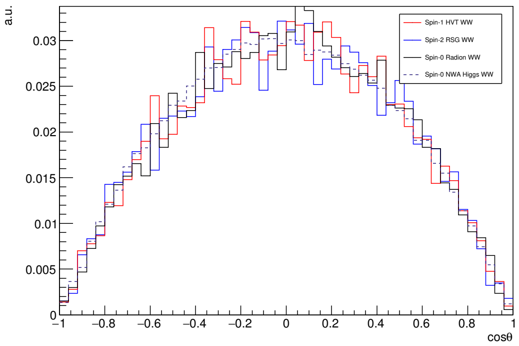


(a) *ggF*

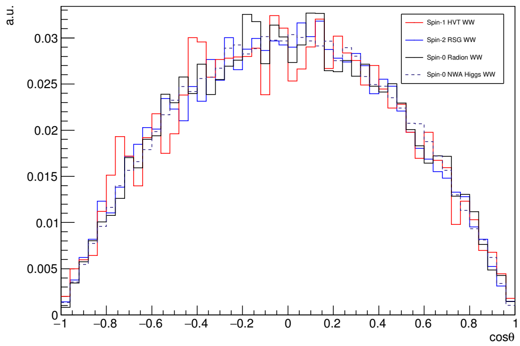


(b) *VBF*

Figure 1.8: Truth mVV distributions for spin-0 models, solid line for Radions and dashed lines for the Heavy Higgs.

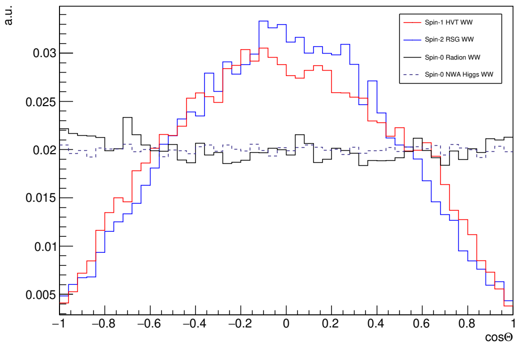


(a) *ggF/DY*

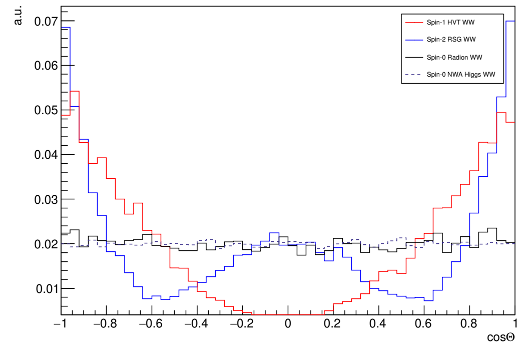


(b) *VBF*

Figure 1.9: Truth $cos\theta$ distribution.

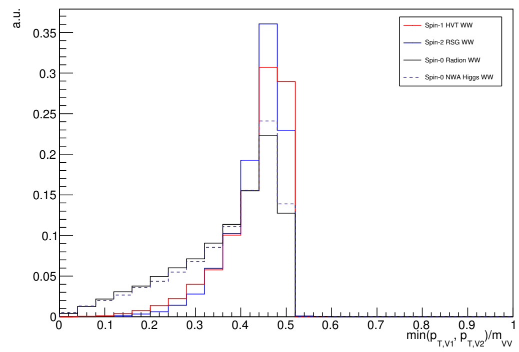


(a) *ggF/DY*

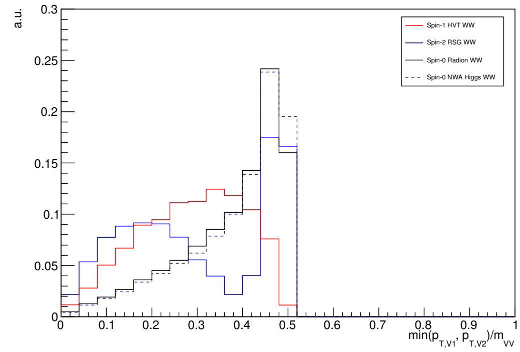


(b) *VBF*

Figure 1.10: Truth $cos\theta^{*}$ in the rest of frame of the resonance.



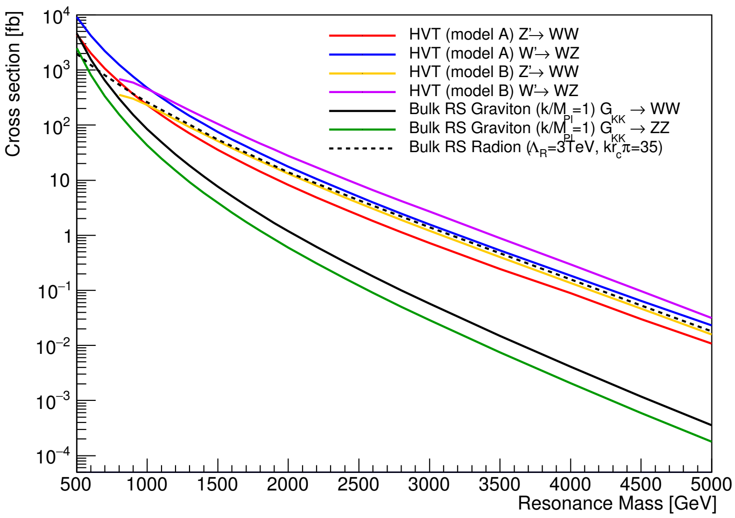
(a) *ggF/DY*



(b) *VBF*

Figure 1.11: Truth p_T/M distributions.

Figure 1.12 shows the expected cross sections, the decay Branching Ratio, and the relative width over the mass for the several benchmark signals and for a couple of mass hypotheses. As examples, Table 1.2 shows the theoretical cross sections, the diboson decay branching ratios, and the total widths of the resonances for two different mass values.



(a)

Figure 1.12: Cross section of the signal models used as benchmark in the resonant analysis.

Model			Spin	$m = 800\ GeV$			$m = 3\ TeV$		
				σ [pb]	BR	Γ/m	σ [fb]	BR	Γ/m
Radion ($k\pi r_c = 35$, $\Lambda_R = 3\ TeV$)		$R \to WW$ $R \to ZZ$	0	0.54 (1.1×10^{-3}) in ggF (VBF)	0.43 0.21	2.6×10^{-3}	1.38 (5.5×10^{-3}) in ggF (VBF)	0.44 0.22	0.032
HVT	Model A	$W\prime \to WZ$ $Z\prime \to WW$	1	53 26	0.024 0.023	0.026	79 36	0.020	0.025
HVT	Model B	$W\prime \to WZ$ $Z\prime \to WW$	1	1.6 0.86	0.43 0.41	0.040	5.5 2.5	0.47	0.031
HVT	Model C (VBF)	$W\prime \to WZ$ $Z\prime \to WW$	1	4.0×10^{-3} 2.7×10^{-3}	0.50 0.49	3.5×10^{-3}	1.6×10^{-3} 1.0×10^{-3}	0.50	3.3×10^{-3}
Bulk RS G_{KK} ($k/\overline{M}_{\text{Planck}} = 1.0$)		$G_{KK} \to WW$ $G_{KK} \to ZZ$	2	1.9 (ggF) 0.050 (VBF)	0.28 0.14	0.037	0.47 (ggF) 1.6×10^{-2} (VBF)	0.20 0.10	0.062

Table 1.2: List of benchmark signal models. Predictions of cross-section σ, branching ratio BR into WW, WZ, or ZZ, and intrinsic width divided by the resonance mass Γ/m, for the given hypothetical new particle at $m = 800\ GeV$ and $3\ TeV$ are summarised.

1.6 Diboson processes and VBF/VBS production mechanism

The final goal of this thesis has been the improvement of the sensitivity of new physics searches in di-boson final states. In particular, the di-boson system is looked for in the semi-leptonic decay channels, in which one boson decays leptonically and the other one hadronically. The decays that have been studied in this analysis are:

- $Z \longrightarrow \nu\nu$,
- $W \longrightarrow l\nu$,
- $Z \longrightarrow ll$, with $l = electron$ or $muon$
- $Z \longrightarrow q\bar{q}$,
- $W \longrightarrow qq'$.

The decay branching ratios of the two weak boson considered are reported in table 1.3.

channel	Branching ratio
$Z \to ll$	$10,1\%$
$Z \to \nu\nu$	$20,0\%$
$Z \to q\bar{q}$	$69,9\%$
$W \to l\nu$	$32,4\%$
$W \to qq'$	$67,6\%$

Table 1.3: Branching ratios of the SM Weak bosons.

According to all the possible decay channels of the Z, W bosons, several di-boson final states are foreseen:

- 0 leptons channel: $ZZ \longrightarrow \nu\nu q\bar{q}$ and $ZW \longrightarrow \nu\nu qq'$;
- 1 lepton channel: $WZ \longrightarrow l\nu q\bar{q}$ and $WW \longrightarrow l\nu qq'$
- 2 leptons channel: $ZZ \longrightarrow llq\bar{q}$ and $ZW \longrightarrow llqq'$

the global branching ratio are reported in the table 1.4. My work involved all the possible channel even if the most focus has been devoted to the 2 leptons channel.

	$Z \to ll$	$Z \to \nu\nu$	$Z \to q\bar{q}$	$W \to l\nu$	$W \to qq'$
$Z \to ll$	$1,02\%$	$2,02\%$	$7,07\%$	$3,28\%$	$6,83\%$
$Z \to \nu\nu$	-	$4,0\%$	$13,9\%$	$6,5\%$	$13,5\%$
$Z \to q\bar{q}$	-	-	$48,7\%$	$22,6\%$	$47,3\%$
$W \to l\nu$	-	-	-	$10,5\%$	$21,9\%$
$W \to qq'$	-	-	-	-	$45,7\%$

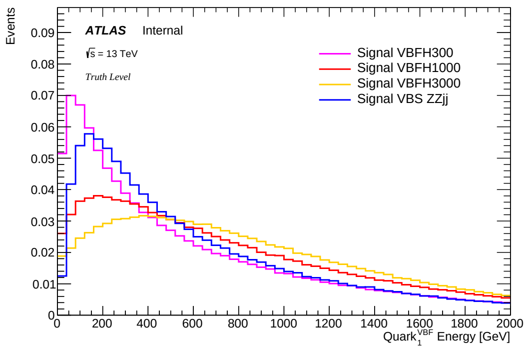
Table 1.4: Branching ratios of the SM di-boson final states.

The semi-leptonic decay channels offer a good environment in which signals of new physics could be found:

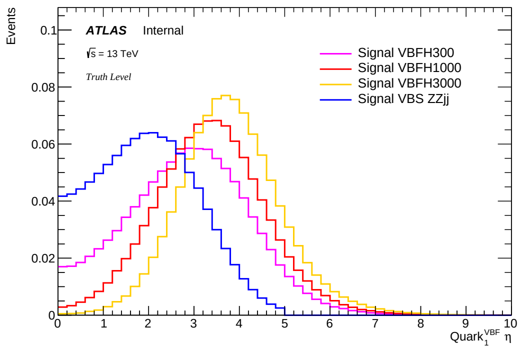
- the requirement of leptons in the final state allows to clean the background environment, especially reducing completely the QCD background;
- the hadronic V decays increases the Branching Ratio (BR) of the entire final state of a factor around 10 respect a search with full leptonic final state (Table 1.4).

All the three leptons channels have been considered in this thesis but the main attention has been focused on the 2 leptons channel. The different di-boson system could allow different spin hypotheses of the new resonce model, as discussed in Section 1.5.1.

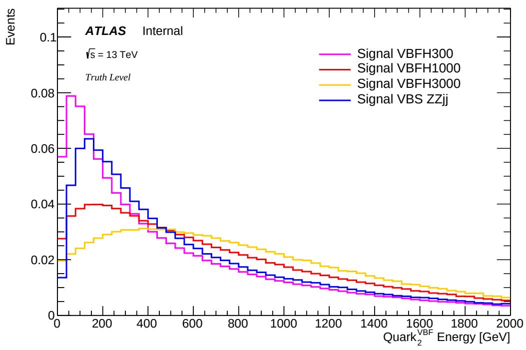
One important point of this work has been the study and the identification of a particular experimental signature of two processes involving di-boson production: *Vector Boson Fusion* and *Vector Boson Scattering*, introduced respectively in Sections 1.4.4-1.5.1. The first process foresees the production of a boson particle trough the fusion of two weak boson scattered from two initial quark coming from the *pp* initial state. The second one foresees the production of two boson not trough a new particle but trough the scattering of two weak boson (Figure 1.7). Both the two processes are characterised by the presence of two additional quark in the final state that means the presence of two extra jets in experimental signature. Figures 1.13 show the kinematic distribution of the two truth quark coming from the VBF/VBS topology. Figures 1.13 show the kinematic distributions of the di-quarks system, built using the two VBF/VBS quarks.



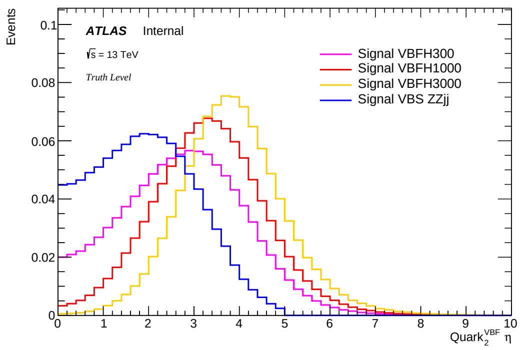
(a)



(b)



(c)



(d)

Figure 1.13: η and energy distributions of the two VBF/VBS quarks for the VBS signal (blue) and for three signal mass for the VBF signal, 300 GeV (magenta), 1000 GeV (red) and 3000 GeV (orange).

The characteristics of these quarks are:

- produced in the forward region, $\eta \sim 2-3$ for the VBS and $\eta \sim 3-5$ for the VBF case depending on the signal mass of the produced resonance.

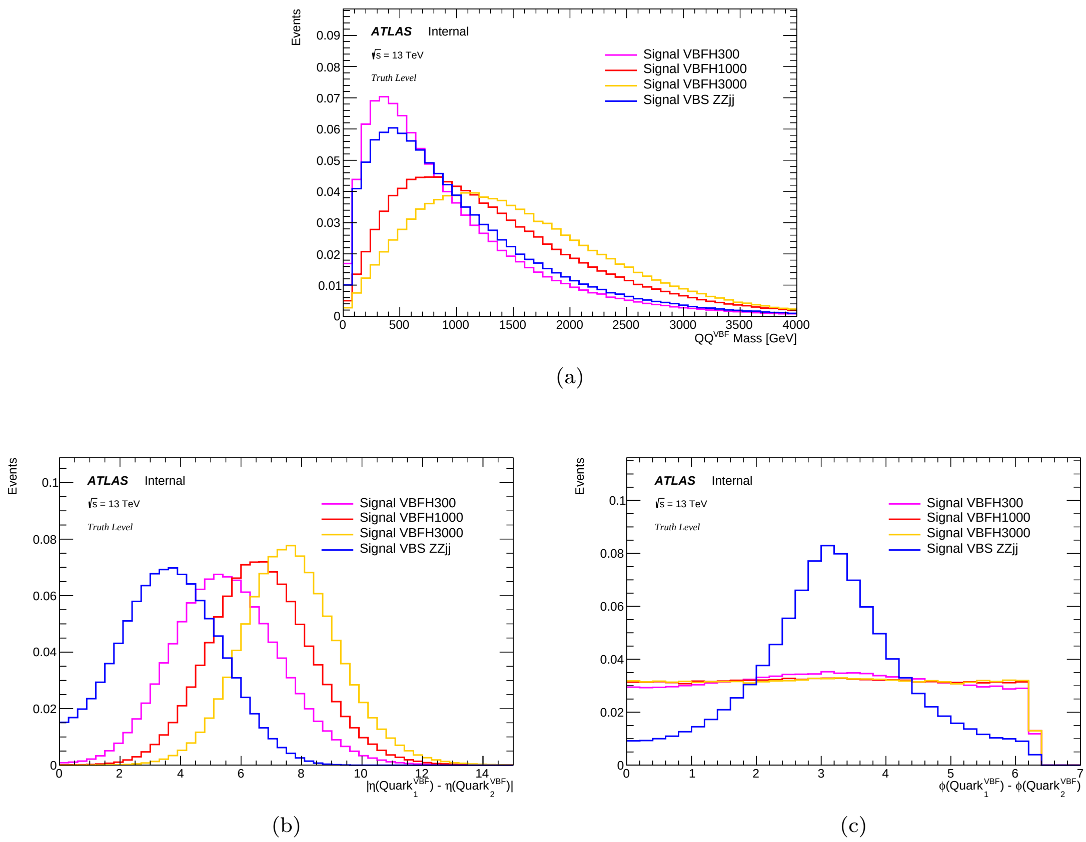
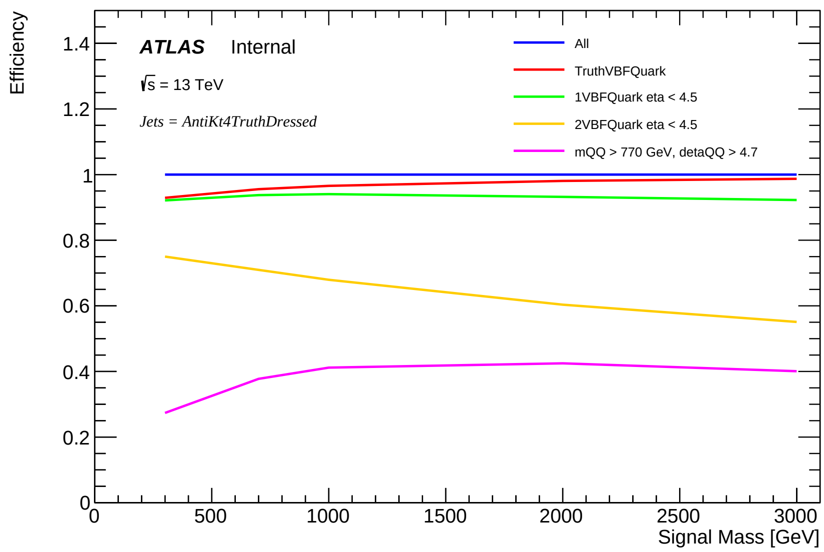


Figure 1.14: Invariant mass (a), η separation and ϕ separation of the di-quark system for the VBS signal (blue) and for three signal mass for the VBF signal, 300 GeV (magenta), 1000 GeV (red) and 3000 GeV (orange).

- the pair of the quarks have high angular separation and have high invariant mass.

These features are expected to be found also in the jets coming from those quarks. In general these features are used to identify these specific topology. As the η distributions show (Figures 1.13) in the case of the VBF production the single quark could have η greater than the ATLAS acceptance (Section 2.2.4) while the VBS case has for $\sim 99\%$ of the case both of the two quark produced inside the ATLAS detector.

Figure 1.15 shows the acceptance of the VBF process, varying the mass of the produced resonance, in the case of one VBF quark inside the ATLAS detector (green curve), in the case in which both the two VBF quark are in the ATLAS acceptance (orange) and, finally, in the case a cut on the angular separation and invariant mass are applied (magenta), these particular value are taken from the previous ATLAS publication for which a cut-based selection on the variables are applied [53].



(a)

Figure 1.15: Truth acceptance of the two VBF quarks as a function of the resonance mass.

Chapter 2

The LHC accelerator and the ATLAS experiment

Founded in 1954, the CERN laboratory is located on the Franco-Swiss border near Geneva. It was one of Europe's first joint ventures and now has 20 member states. The name CERN is derived from the acronym for the French *Conseil Européen pour la Recherche Nucléaire*, or European Council for Nuclear Research, a provisional body founded in 1952 with the mandate of establishing a world-class fundamental physics research organization in Europe. At that time, pure physics research concentrated on understanding the inside of the atom, hence the word "nuclear". At this moment, the understanding of matter goes much deeper than the nucleus, and the main area of research at CERN is particle physics − the study of the fundamental constituents of matter and the forces acting between them.

At CERN, the European Organization for Nuclear Research, physicists and engineers are probing the fundamental structure of the universe. They use the world's largest and most complex particle accelerator to study the basic constituents of matter or the fundamental particles. The particles are made to collide together at close to the speed of light. The process gives the physicists clues about how the particles interact, and provides insights into the fundamental laws of nature.

The instruments used at CERN are purpose-built particle accelerators and detectors. Accelerators boost beams of particles to high energies before the beams are made to collide with each other. Detectors −like ATLAS− observe and record the results of these collisions.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC), conceptualised around a quarter of century back, is built in a circular tunnel 27 km in circumference. The tunnel is buried around 50 m to 175 m underground. It located between the Swiss and French borders.

It is a proton-proton (pp) collider, and the collision were delivered at $\sqrt{s} = 7 - 8\,TeV$ during the Run-I (2010-2012) while they are being collected at $\sqrt{s} = 13\,TeV$ during Run-II (2015-2018).
The first beams were circulated successfully on 10^{th} *September* 2008. Unfortunately on 19^{th} *September* a serious fault developed damaging a number of superconducting magnets. The repair required a long technical intervention. The LHC beam did not see beam again before $November$ 2009. First collisions took place on 30^{th} $March$ 2010 with the rest of the year mainly devoted to commissioning. The 2011 was the first production year with $5\,fb^{-1}$ delivered to both ATLAS and CMS with a center mass energy of $\sqrt{7}\ TeV$. The 2012 started with over $6\,fb^{-1}$ delivered by the time of the summer conferences and it was ended with around 20 fb^{-1} and with a center mass energy of $\sqrt{8}\ TeV$.

Data that allowed for the announcement of the discovery of a Higgs-like particle on 4^{th} $July$ 2012 mentioned in the previous chapter (Section 1.3.1).

During the 2015 a second phase of event production at LHC, called *Run-II*, started. During the Run-II the LHC reached an energy collision of $13\,TeV$ and the purpose of the program was to collect data corresponding to more than $100\,fb^{-1}$ and it has been reached with the 140 fb^{-1} collected during $2015-2018$. The Run-II started in May 2015. An initial phase of collision with $50\,ns$ bunch spacing and $1\,fb^{-1}$ in luminosity took place; the data collected were dedicated to control the performances of the magnet and of the alignment of the spectrometer.

Just after the commitioning phase the beams, characterized by $25\,ns$ bunch-spacing, circulated in the accelerator and produced collisions at $\sqrt{s} = 13\,TeV$ with a peak luminosity of $5,0 \cdot 10^{33} cm^{-2}s^{-1}$.

One of the crucial parameters for the discovery power of a particle collider is the *instantaneous luminosity*, $\mathcal{L}$, since it is proportional to the event rate $\frac{dN}{dt}$:

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma \tag{2.1}$$

where σ is the cross section of the considered process. The instantaneous luminosity of a particle accelerator depends on its intrinsic features:

$$\mathcal{L} = \frac{N_p^2 f k}{4\pi R^2} \tag{2.2}$$

where N_p is the number of protons in each bunch, f is the revolution frequency of the protons in the accelerating ring, k is the number of bunches circulating in the beam and R is the mean radius of the proton distribution on the plane orthogonal to the beam direction.

The instantaneous luminosity delivered by the LHC reached the value of $21.0 \cdot 10^{33}$ $cm^{-2}s^{-1}$ at its maximum (Figure 2.1), where the design peak luminosity was 10^{34} $cm^{-2}s^{-1}$. This high luminosity is reached with 2220 (2808 from the design) bunches per beam, each of them containing 10^{11} protons. The bunches have very small transverse spread, about $15\,\mu m$ in the transverse direction, and the longitudinal length is about $???\,cm$. As in the design of the LHC the bunches crossed every $25\,ns$ during the run-II, giving a collision rate of $40\,MHz$. This setting implies a mean number of interaction per bunch crossing between $20-50$ as showed in Figure 2.1. These parameters achieved allowed an integrated luminosity showed in Figure 2.2.

Feature	design value	actual value
beam energy $[TeV]$	7	$6,5$
bunch spacing $[ns]$	25	25
peak luminosity $[cm^{-2}s^{-1}]$	$10^{34}cm^{-2}s^{-1}$	$12,1 \cdot 10^{33}cm^{-2}s^{-1}$
mean number of interaction per bunch crossing	19	33
number of bunches	2808	2220
protons per bunch	$1,15 \cdot 10^{11}$	$1,18 \cdot 10^{11}$
bunch transverse dimensions $[\mu m]$	$16\,\mu m$	$4\,\mu m$

Table 2.1: Capabilities of the LHC during the Run-II. The first column contains the values as in the LHC design, the second column contains the actual value of the features.

Part of the acceleration chain and the different positions of the LHC's experiments are showed in Figure 2.3: after their production and an of $1.4\,GeV$, the Super Proton Synchrotron raises their energy up to $450\,GeV$ before injecting them into the LHC. Ones there, the protons are accelerated in the two opposite directions up to the colliding energy of $3.5\,TeV$ (2011), $4\,TeV$ (2012) and $6,5$ TeV (2015/18) per beam.

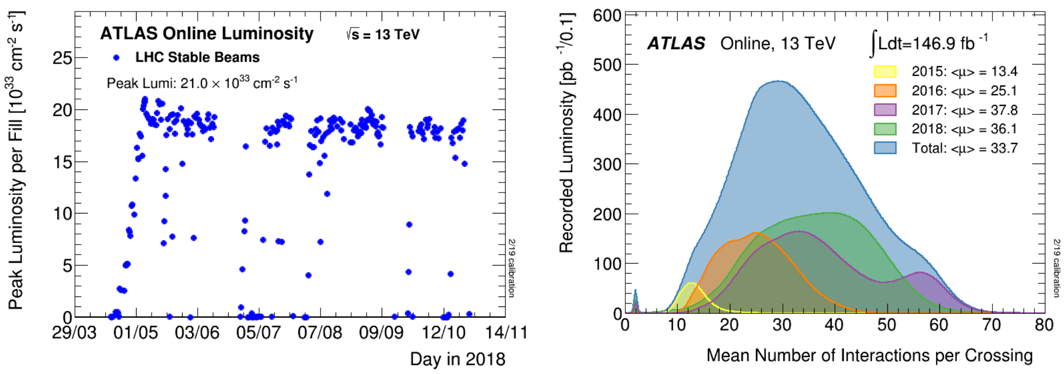


Figure 2.1: The LHC peak instantaneous luminosity during 2018 year of run-II (a) and Distribution of the mean number of interactions per crossing for the 2018 pp collision data at 13 TeV centre-of-mass energy (b).

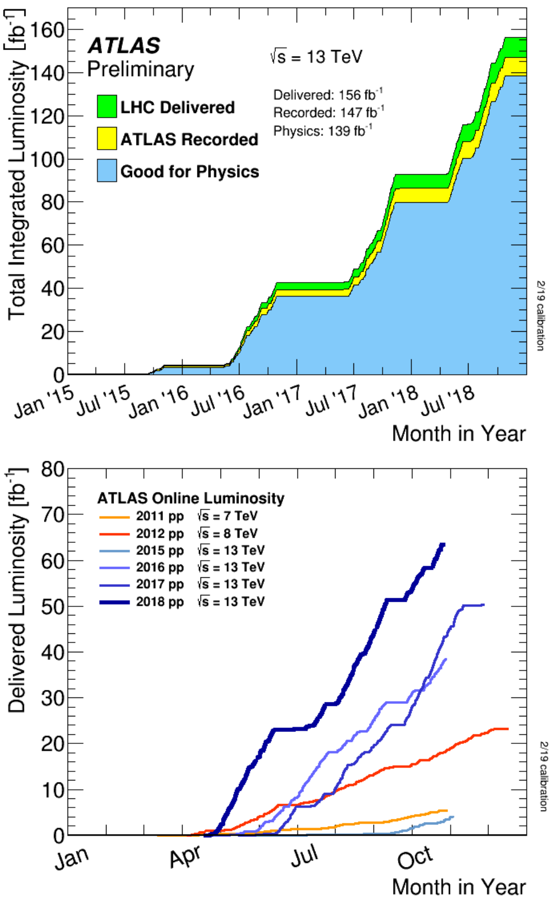


Figure 2.2: The integrated luminosity as a function of time delivered by LHC (green), recorded by ATLAS (yellow) and good for physics (azure) during Run-II (a). Comparison of integrated luminosity of LHC during the different years of operations of run-I and run-II (b).

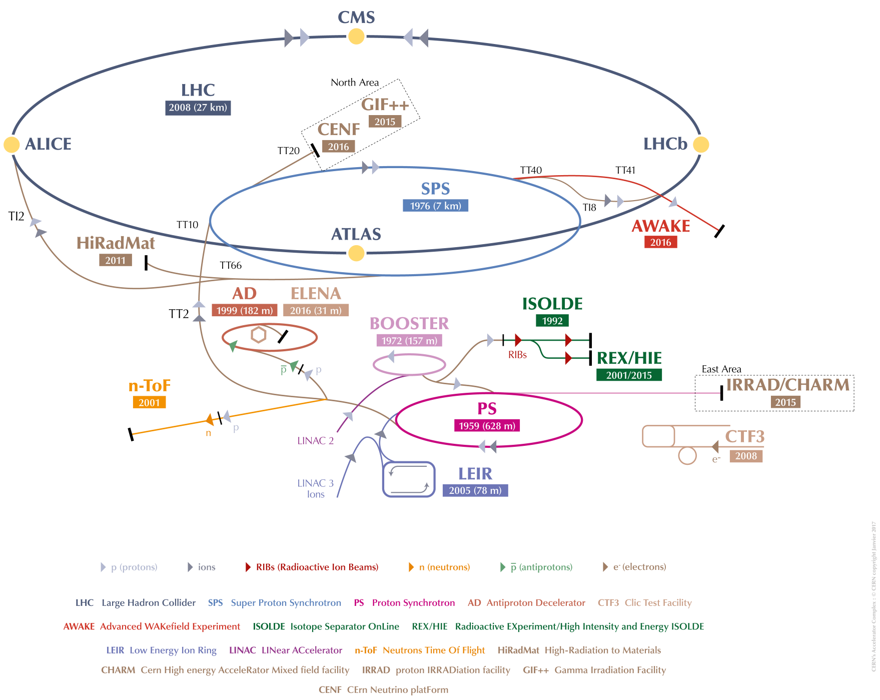


Figure 2.3: The LHC particles accelerator, in which it is possible to see the SPS and the different beam's collision points with their corresponding experiment.

Since LHC accelerates two beams of same sign particles, two separate accelerating cavities and two different magnetic fields are needed: LHC is equipped with 1232 superconducting magnets and 16 radiofrequency cavities which bend and accelerate the proton beams in the two parallel beam lines in the machine. The magnetic field used to bend such energetic proton beams is of $8.3\,T$ and to reach such a magnetic fields the superconducting magnets are cooled down to $1.9\,K$ and a $13\,kA$ current circulates inside them.

The LHC provides collisions in four collision points along its circumference where detector experiments located:

- ALICE (*A Large Ion ColliderExperiment*),
- ATLAS (*A Toroidal Lhc ApparatuS*),
- CMS (*Compact Muon Solenoid*),
- LHCb (*Large Hadron Collider beauty*).

ATLAS and CMS are multi-purpose detectors, while ALICE and LHCb are focused on more specific studies: (See Figure 2.3) ALICE focuses on the quark-gluon plasma produced in heavy-ions collisions[1], while LHCb focuses on the study of CP violation processes occurring in b and c hadron decays.

The total integrated luminosity as a function of the days of running during the run-II are represented in Figures 2.2; at the end of the run-II, the integrated luminosity is of 3.9 fb^{-1} for 2015, 35.6 fb^{-1} for 2016, 46.9 fb^{-1} for 2017 and 60.6 fb^{-1} for 2018. The total integrated luminosity of the run-II data taking overcame the luminosity reached in the run-I of a factor ~ 7 (Figure 2.2).

[1]The LHC is able to accelerate and collide lead ions at $\sqrt{s} = 2.76\,TeV$ per nucleon, and ions collisions are foreseen each year in the LHC program. Not part of our actual studies.

2.2 The ATLAS Detector

The ATLAS detector is one of the four main experiments recording the collisions provided by the LHC. It is 20 m tall and 45 m long and weights more than 7000 $tons$.

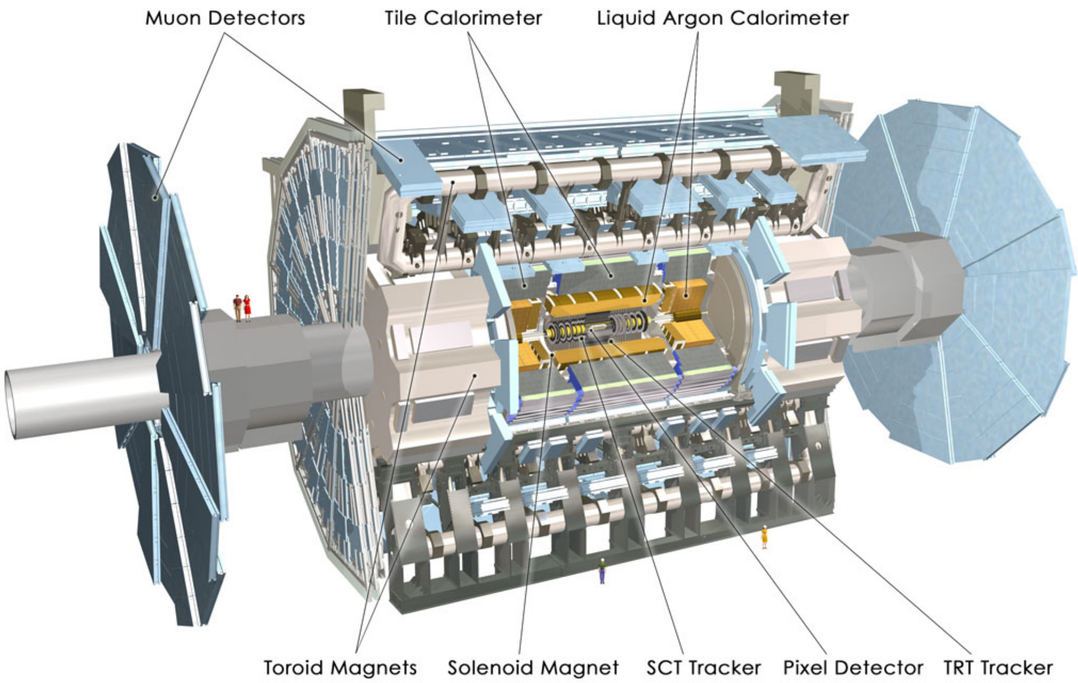


Figure 2.4: The ATLAS Detector: all the subdetectors it is composed of are shown.

The structure has a cylindrical shape centred at the interaction point with its axis along the beam line, and it is composed of several concentric subdeterctors which measure different features of the particles generated in the *pp* collision as they fly from the center of the detector to the outer part, as shown in Figure 2.5. From the innermost to the outermost layer, the ATLAS experiment is composed of (see Figure 2.4):

- An inner tracking system to detect charged particles and measure their momentum and direction.
- A solenoidal superconducting magnet providing a uniform magnetic field along the beam axis in which the inner detector is immersed.
- An electromagnetic calorimeter to measure the energy deposited by electrons and photons.
- An hadronic calorimeter to measure the energy deposited by hadrons.
- A muon spectrometer, that is a tracking system for the measurement of muons as they travel throughout all the detector and are the only particles reaching the outer part.
- An air-cored superconducting toroidal magnet system which provide the magnetic field to the muon spectrometer.

In the following sections details about the structure of the subdetectors are be given, as well as some insight about how they work and their performances.

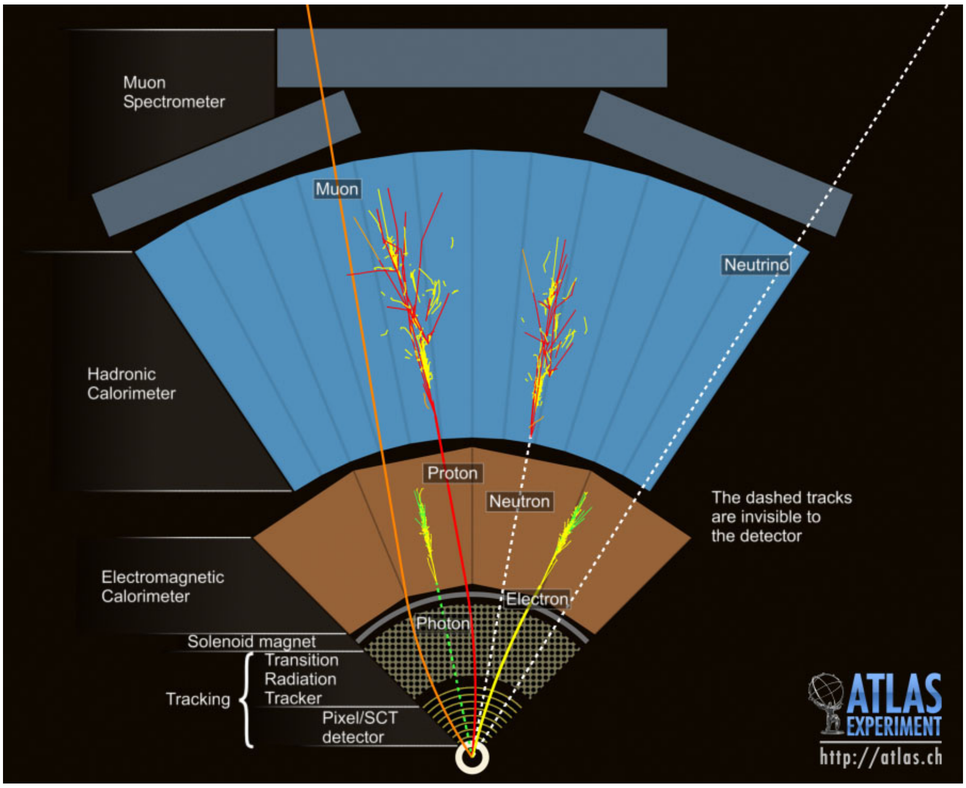


Figure 2.5: Schema of the detection of the particles produced in a proton collision while they travel through the several layers of the ATLAS detector.

2.2.1 ATLAS Coordinate System

The ATLAS coordinate system is a cartesian right-handed coordinate system, with the nominal collision point as the origin (Figure 2.6), the z axis is along the beam line and the $x - y$ plane is the plane perpendicular to the beam line.

The x axis points to the center of the LHC ring, while the y axis goes upwards. The azimuthal angle ϕ is defined around the beam axis, while the polar angle θ is the angle from the z axis in the $y - z$ plane. The θ variable is not invariant under boosts along the z axis, and so instead of the θ angle the *pseudorapidity* [2] η is used:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{2.3}$$

Since at an hadronic collider the real colliding particles are the partons inside the protons, we can say that the actual center of mass energy in unknown in each collision:

$$\hat{s} = x_1 \cdot x_2 \cdot s \tag{2.4}$$

where $\hat{s}$ is the effective collision energy, x_1 and x_2 are the fractions of momentum carried by the two colliding partons and s is the colliding energy of the two protons. Because of this, the total momentum along the beam axis before the collision is unknown, while the total momentum in the transverse plane (i.e. the $x - y$ plane) is known to be zero and hence we can apply the momentum and energy conservation laws only on the transverse plane (because we know what is the initial

[2] Actually the real boost-invariant variable is the *rapidity* y: $y = \frac{1}{2}\ln\frac{E+p\cos\theta}{E-p\cos\theta}$. In the ultra-relativistic limit the rapidity y can be substituted with the pseudorapidity η.

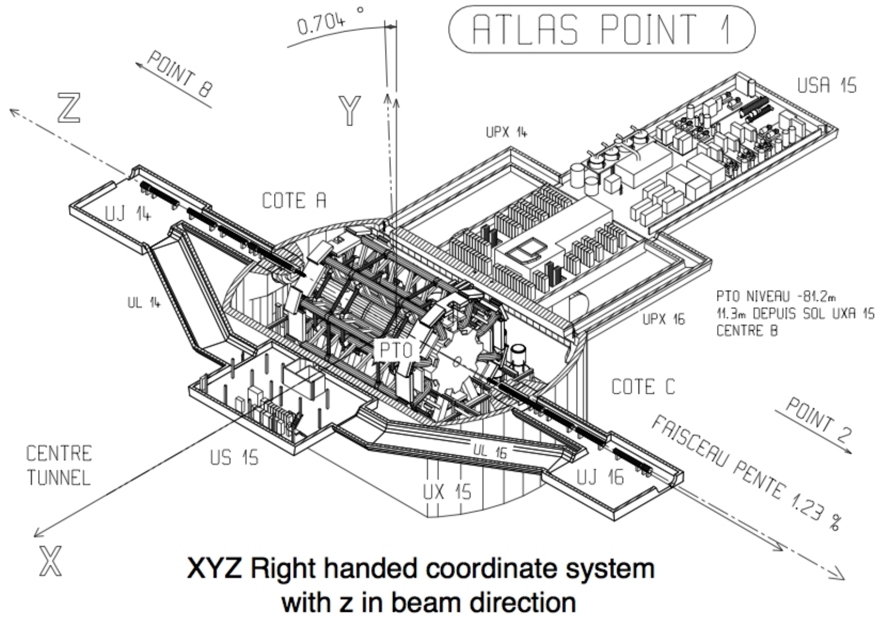


Figure 2.6: Reference system used in ATLAS.

total momentum).

For this reason transverse quantities are considered, and they will be denoted with the "T" subscript (e.g. p_T stands for transverse momentum, that is the projection of the momentum on the $x-y$ plane).

2.2.2 ATLAS Magnets

The ATLAS detector is equipped with two magnetic systems: a superconducting solenoid [54], providing a magnetic field to the inner tracking system, and a system of air-core superconducting toroidal magnets [55, 56] located in the outer part of the detector as shown in Figure 2.7.

The solenoid covers the central region region of the detector, provides an uniform magnetic field of approximately $2\,T$ along the z axis bending tracks of the particles in the transverse plane in order to let the inner tracking system measure their transverse momentum.

The solenoid is located between the inner detector and the electromagnetic calorimeter and its dimensions (its width, particularly) have been optimized in order to minimize the amount of dead material (only 0.83 radiation lengths) in front of the calorimetric system.

The toroid is one of the peculiarities of the ATLAS detector: it is located outside of the calorimetric system covers the region $|\eta| < 3$ (considering all its subparts), and provides a magnetic field whose peak intensities are $3.9\,T$ in the central region of the detector and $4.1\,T$ in the forward region.

The aim of such a toroid is to have a large lever arm to improve the measurement of the muon transverse momentum, and it is built "in air" in order to minimize the muon multiple scattering within the detector.

The ATLAS double magnetic system has been designed to provide two independent measurements of the muon transverse momentum in the inner detector and in the muon spectrometer, thus ensuring good muon momentum resolution from few GeV up to the TeV scale.

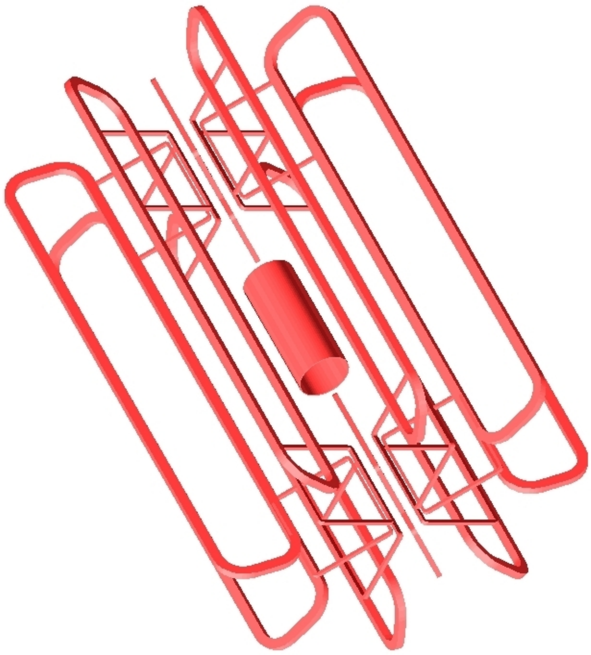


Figure 2.7: The magnetic system of the ATLAS detector: the inner cylinder is the superconducting solenoid, while the external parts are the coils of the toroid.

2.2.3 The Inner Detector

The ATLAS Inner Detector tracker (ID), shown in Figure 2.8, is composed by three concentric cylindrical subdetectors. Its axis is centred on the z axis and it is approximately 6 m long and its diameter $2.30\,m$, covering the region ($|\eta| < 2.5$).

The three sub-detector into the ID are:

- **Pixel Detector:** it is composed of three layers of silicon pixels. It provides high-precision track measurement, since the spatial resolution on the single hit is ~ 10 μm in the ϕ coordinate and ~ 115 μm along the z coordinate.
- **Semiconductor Tracker (SCT):** it is the second high-precision detector of the ATLAS inner tracker. It is composed of eight layers of silicon strips with a spatial resolution on the single hit of 17 μm in ϕ and 580 μm along z. The Pixel Detector and the Semiconductor Tracker together provide on average eight high-precision hits per track.
- **Trasition Radiation Tracker (TRT):** it is composed of straw tubes chambers. The resolution of such a detector is lower than the previous one (~ 130 μm per straw), but it is compensated by the high number of points per track (36 on average) that it can provide.

The aim of the ATLAS ID is to measure the tracks of the charged particles produced in the pp collision and all the related features: p_T, η, ϕ, the eventual secondary vertexes due to long-lived particles.

The momentum is measured by measuring the track curvature in the magnetic field provided by the superconducting solenoid described in the previous section. To estimate the expected resolution the *sagitta method* can be used: the magnetic field bends the trajectory of the charged

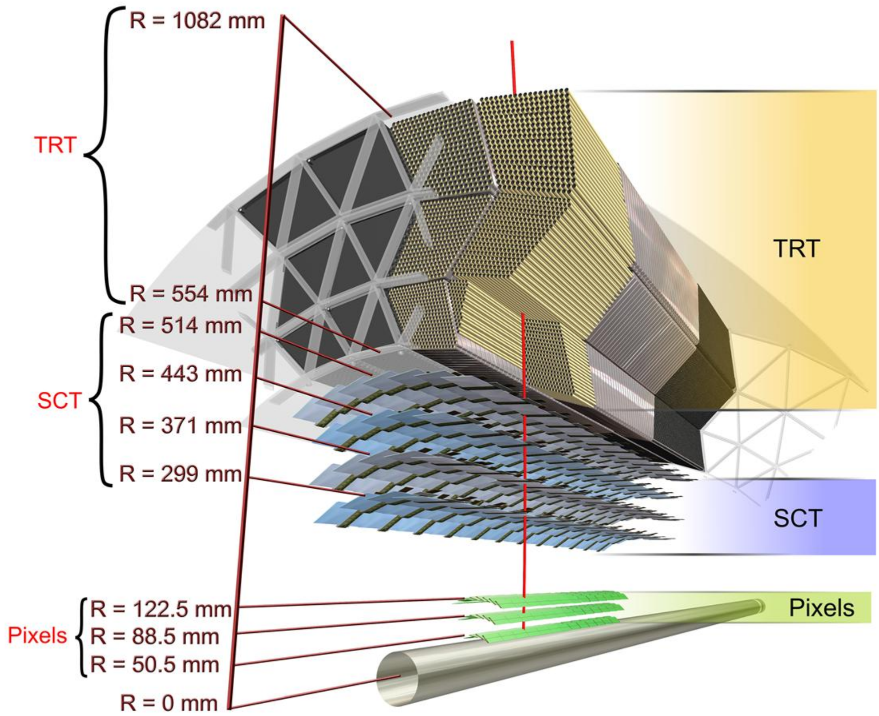


Figure 2.8: The ATLAS Inner Detector tracker: the three subdetectors (the Pixel Detector, the Semiconductor Tracker and the Transition Radiation Tracker) are shown as well as their radial dimensions.

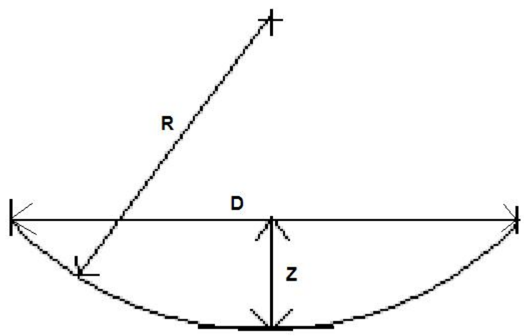


Figure 2.9: The sagitta of a track is the maximum distance between the track itself (that is an arc of a circle) and the straight segment having the same starting and ending points.

particles in the ϕ coordinate because of Lorentz's force:

$$\vec{F}_L = q\vec{v} \times \vec{B} \tag{2.5}$$

where q is the charge of the particle, $\vec{v}$ is its velocity and $\vec{B}$ is the magnetic field. The resolution of the momentum measurement depends on many detector-related parameters:

$$\frac{\Delta p}{p^2} = \frac{8}{0.3 \cdot B \cdot L^2} \Delta s \tag{2.6}$$

where B is the magnetic field expressed in Tesla, L is the lenght of the reconstructed track expressed in meters, while Δs is (see Figure 2.9):

$$\Delta s = \frac{\epsilon}{8} \sqrt{\frac{720}{N+4}} \tag{2.7}$$

where N is the number of measured points on the track and ϵ is the resolution on the measurement of the points.

From formulas 2.7 and 2.6 it is possible to see how it is crucial to have a strong magnetic filed, an high number of points per track and a good spatial resolution on these points in order to have a good resolution on the track p_T .

2.2.4 The Calorimetric System

In an high-energy physics experiment the calorimeters are used to measure the energy of photons, electrons (the electromagnetic calorimeter), hadronic jets (hadronic calorimeter) and the missing E_T (due to undetected particles like neutrinos) which is measured thanks to the tightness of the calorimetric system.

The ATLAS calorimeter has a cylindrical shape centered around the interaction point with its axis lying on the ATLAS z axis. It is long about $13\,m$ and the external radii of the electromagnetic and hadronic calorimeters are 2.25 and $4.25\,m$ respectively.

The ATLAS calorimeters are represented in Figure 2.10 and the absorption lengths as a function of η are shown in Figure 2.11.

The Electromagnetic Calorimeter

The Electromagnetic Calorimeter of the ATLAS experiment covers the region up to $|\eta| < 3.2$. The structure of the Electromagnetic Calorimeter has a special feature, how you can see in Figure 2.12: it has an accordion structure made of lead (whose thickness varies as a function of η in order to maximise the energy resolution) which is immersed in liquid Argon, which is the active material of the calorimeter. This structure confers to the calorimeter very high acceptance and symmetry in the ϕ coordinate.

In the central region $|\eta| < 2.5$ the radial coordinate the electromagnetic calorimeter has three sampling channels in order to maximize particle identification power (see Figure 2.12).

The calorimeter is segmented in cells of variable dimensions as a function of η as well as its thickness ($> 24X_0$ in the central region and $> 26X_0$ in the forward region): in the central region the segmentation is $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$.

The ATLAS EM calorimeter energy resolution is parametrized as:

$$\frac{\Delta E}{E} = \frac{10\%}{\sqrt{E[GeV]}} \oplus 1\% \tag{2.8}$$

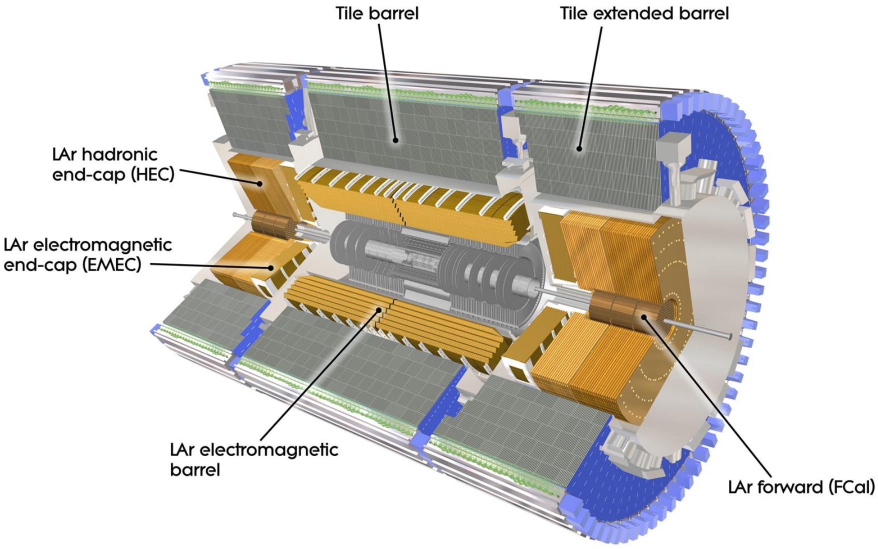


Figure 2.10: The ATLAS calorimetric system: the electromagnetic calorimeter made of liquid Argon and Lead and the hadronic caloimeter, whose composition varies as a function of η.

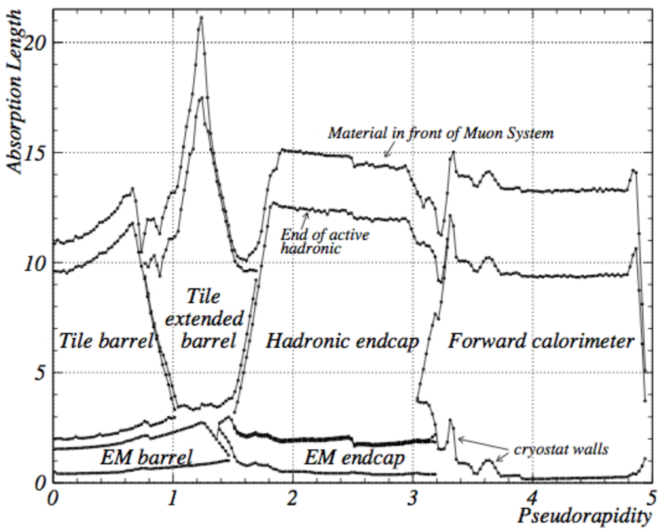


Figure 2.11: Amount of material in terms of absorption length in the ATLAS calorimetric system as a function of η.

Where 10% is the sampling term and 1% is the constant inter-calibration term. The η resolution is:

$$\frac{40\,mrad}{\sqrt{E[GeV]}}. \tag{2.9}$$

The Hadronic Calorimeter

The Hadronic Calorimeter covers the region $|\eta| < 4.5$, and it is realized with a variety of techniques as a function of η like it is possible to check in Figure 2.10.

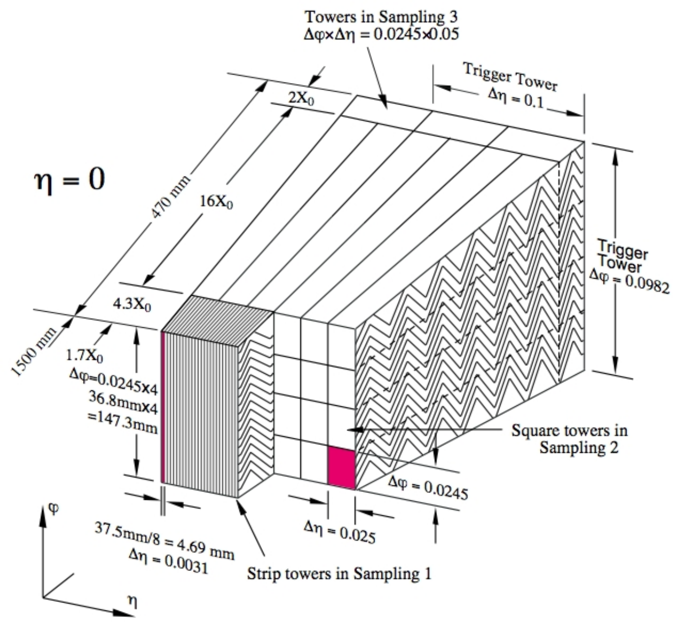


Figure 2.12: The accordion structure of the electromagnetic calorimeter and its radial segmentation.

The central region ($|\eta| < 1.7$) it is made of alternating layers of iron (used as absorber) and scintillating tiles as active material, and its thickness offers about 10 interactions lengths λ at $\eta = 0$. It is segmented in $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ pseudo-projective towers pointing to the interaction point.

The "endcap" region ($1.7 < |\eta| < 3.1$) is equipped with a liquid Argon and lead, as the Electromagnetic Calorimeter, while the forward region ($3.1 < |\eta| < 4.5$) is equipped with liquid Argon, but the accordion structure is replaced by a concentric rods and tubes made of copper. This variety of materials and structures is due to the different radiation hardness required in the different parts of the detector.

2.2.5 The Muon Spectrometer

The ATLAS Muon Spectrometer (MS) is instrumented with both trigger and high-precision chambers immersed in the magnetic field provided by the toroidal magnets which bends the particles along the η coordinate, and it allows to measure the muons p_T in the region $|\eta| < 2.7$ using the sagitta method described in Section 2.2.3. Here the lenght of the lever arm plays a leading role on the p_T resolution. The MS is shown in Figure 2.13.

The chambers used to reconstruct the muon track are of several types depending on the η region, in order to face the different rate conditions present in the different parts of the detector. In the central region ($|\eta| < 2$) Monitored Drift Tubes (MDTs) are used. The MTD chambers are composed of aluminium tubes of 30 mm diameter and 400 μm thickness, with a 50 μm diameter central wire. The tubes are filled with a mixture of Argon and CO_2 at high pressure (3 bars), and each tube has a spatial resolution of 80 μm.

At higher pseudo-rapidity ($2 < |\eta| < 2.7$) the higher granularity of the Cathode Strip Chambers (CSC) are used. CSC chambers are multiwire proportional chambers in which the readout is per-

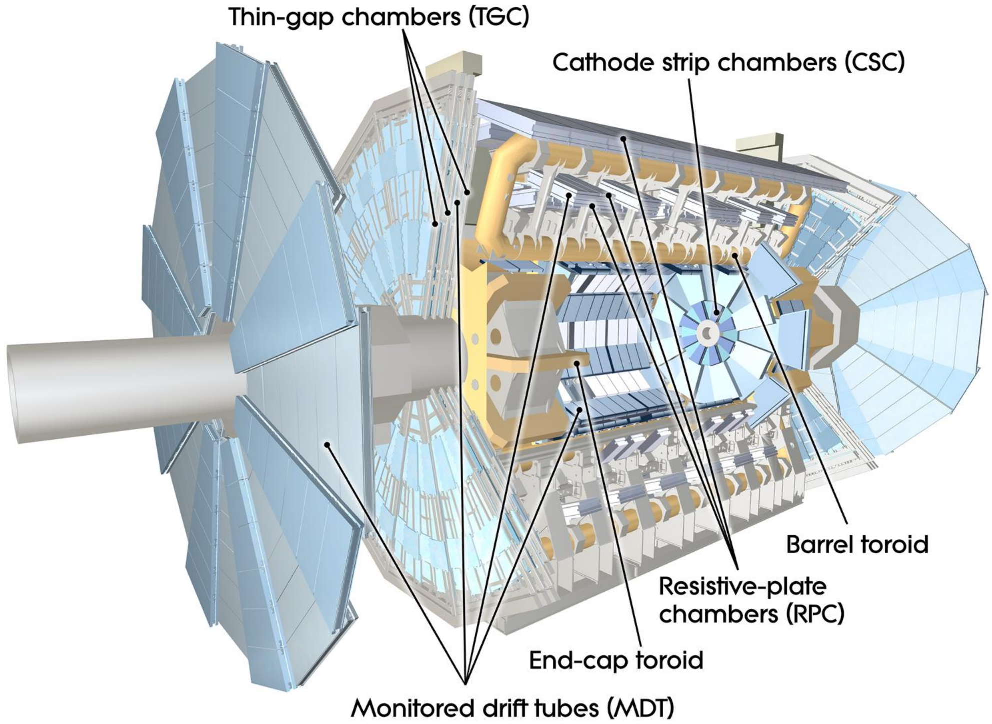


Figure 2.13: The ATLAS Muon Spectrometer.

formed using strips forming a grid on the cathode plane in both orthogonal and parallel direction with respect to the wire. The spatial resolution of the CSC is about 60 μm.

As shown in Figure 2.13, in the central region the MS is arranged on a three layer −or stations− cylindrical structure which radii are 5, 7.5 and 10 m; while in the forward region the detectors are arranged vertically, forming four disks at 7, 10, 14 and $21-23\,m$ from the interaction point.

The other chambers installed on the spectrometer are used for the trigger (see next section for details). The chambers used for the muon trigger are Resistive Plate Chambers (RPC) in the central region ($|\eta| < 1.05$) and Thin Gas Chambers (TGC) in the forward region.

These detectors provide very high time resolution ($\mathcal{O}$(ns)) even if the spatial resolution is not so high ($\mathcal{O}$(cm)).

The spectrometer has been designed to measure the muon p_T up to $1\,TeV$ with an error of less then 10%; this feature was required to optimize the Higgs boson discovery potential.

2.3 The ATLAS Trigger system

The LHC is designed to provide collisions at a frequency of $40\,MHz$ and, since the average dimension of an ATLAS event is $\sim 1.5\,MB$, a recording rate of $\sim 60\,TB$ per second would be needed, while the current technology allows to record data at about 300 MB/s. To deal with this environment and knowing that the interesting physics at LHC occurs at very low rate the events to be recorded can be selected without loosing the relevant information.

This selection is performed on-line by the ATLAS trigger and data acquisition system [57]. The ATLAS trigger is designed to rapidly inspect the events detected by the ATLAS detector and

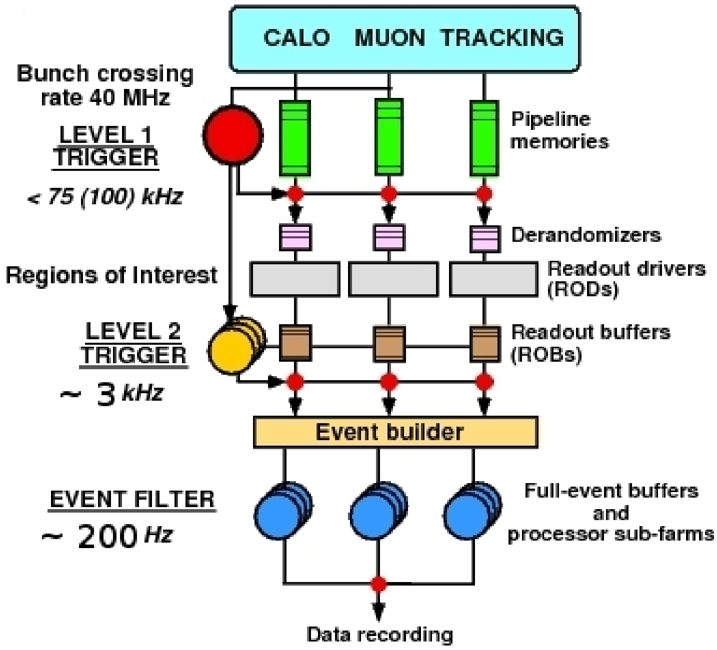


Figure 2.14: Main structure of the ATLAS trigger system: it is made of three levels, each improving the measurement of the previous levels also combining informations from different subdetectors.

choose whether record or discard the event after having compared its main features with a set of predefined thresholds contained in the trigger menu. In case that the trigger decides to discard an event, then the event is not recorded and lost forever.

The ATLAS trigger system has a three level structure: each level refines the measurements of the previous level introducing also new selection criteria and combining the information from different subdetectors, as shown in Figure 2.14.

The first level of the ATLAS trigger (L1 or LVL1) is completely hardware-based and it makes use of only the data collected by the calorimetric system and the muon spectrometer: the L1 trigger only looks for high-p_T muons candidates or calorimetric objects (electrons/γ, jets) by means of fast and rough measurements performed by ad-hoc detectors in the Muon Spectrometer (RPC, TGC) and simplified object identification in the calorimeter.

The L1 is designed to take a decision on the event in 2.5 μs and its output is a list of so-called *Regions of Interest* (RoI), which are $\eta - \phi$ regions of the detector in which interesting activity has been detected, and the output rate is about 100 kHz.

The second level of the ATLAS trigger (L2 or LVL2) is completely software-based. It takes as input the RoIs provided by the L1, and refines the measurement into these regions: data of the precision chambers are used in the Muon Spectrometer (MDT, CSC) as well as the data from the ID, while the measurement of the calorimetric objects is refined using higher level algorithms.

Moreover the data of the different subdetectors are combined together in order to obtained better object reconstruction/identification (e.g. the ID and the MS tracks are combined for the muons, ID and calorimetric informations are combined to discriminate between electrons and photons). The L2 takes its decision in $\mathcal{O}$(10ms) and its output rate is about 3 kHz.

The third level of the ATLAS trigger (Event Filter, EF) is completely software-based and forms, together with the L2, the High Level Trigger (HLT). At this stage a full reconstruction of the detector is performed (the measurement is not restricted to the RoIs), and the algorithms run at the EF are mostly the off-line reconstruction algorithms adapted to the on-line environment. The decision of the EF is taken in $\mathcal{O}(1\text{s})$ and the output rate is about 400 Hz.

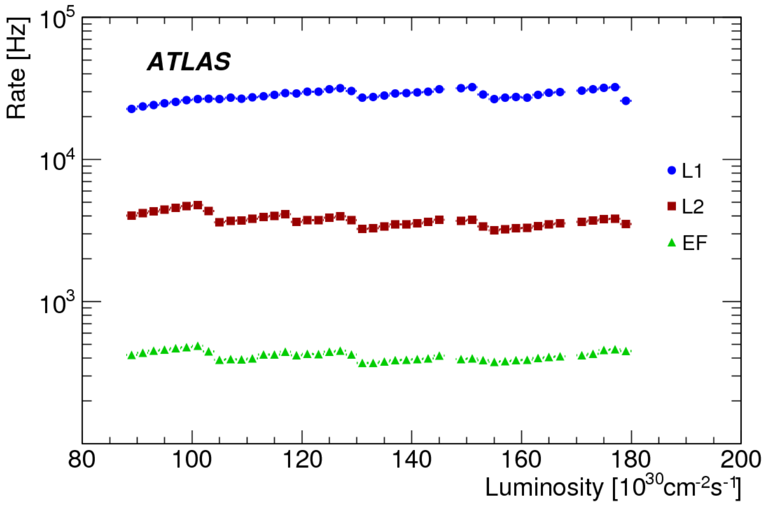


Figure 2.15: Total trigger rates at each level of the ATLAS trigger.

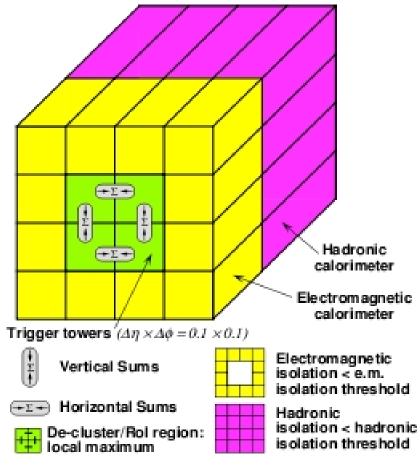


Figure 2.16: The L1 trigger for calorimetric objects in the Electromagnetic Calorimeter: the green area represents the RoI cluster, the yellow area is the region used for the isolation requirements, and the pink area is the region used for the hadronic isolation.

Figure 2.15 shows the total trigger rate for all the three levels as a function of the instantaneous luminosity: how can be seen the trigger rates are kept stable. This happens thanks to changes in the prescales and in the trigger menu −on-line into the ATLAS Control Room−, where higher thresholds or quality criteria on the trigger objects are required as the luminosity increases.

2.3.1 Electron Trigger

The electron trigger follows the three level ATLAS trigger structure, in which the measurements and the selections are refined at each stage. At the first level the electron trigger makes use only of the calorimeters, and hence no distinction between electrons and photons is possible since they are both identified as "calorimetric objects". In particular the L1 trigger measurement is a real calorimetric measurement even if it is done with reduced granularity, represented in Figure 2.16:

Once a relevant amount of energy is detected, the total energy in a little 2×2 cluster is measured (green area), and the isolation with respect to electromagnetic (yellow area) and hadronic activity (pink area, e.g. due to electrons coming from heavy quark decay) is computed. If the these three parameters (E_T, electromagnetic and hadronic isolation) fulfil the requirements, then the electromagnetic calorimeter is accepted as a good calorimetric object and its RoI is propagated to the L2.

The L2 trigger basically refines the calorimetric measurement, accessing the full granularity of the calorimeters and studying the shape of the energy deposit (e.g. π^0/γ separation), and includes the data of the inner tracking system. At this level a "calorimetric object" may become an electron if an ID track consistent with it is found. Since the measurements are more precise at this level, tighter conditions on the quality and the kinematic features of the electron candidates can be required.

At the end of the chain the EF further refines the measurements performed at the L2 on the electron candidates, running algorithms very similar to the off-line ones and having access to the data of all the subdetectors with full granularity.

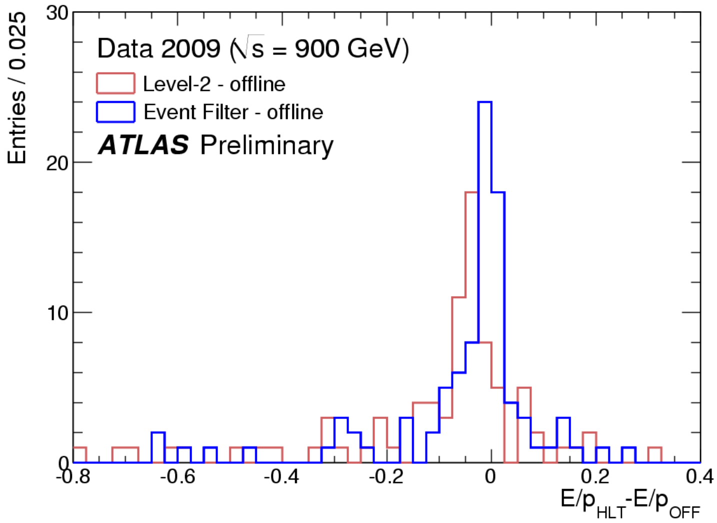


Figure 2.17: E/p distribution found by the HLT and the off-line for the electron trigger. The distributions are shown for L2 and EF separately.

The distribution of the difference between the off-line and the value measured at different trigger levels of the E/p variables for electrons is showed in Figure 2.17. This shows how the EF measurement (blue line) is better than the L2 measurement (red line), since the former is allowed to use reconstruction algorithms very similar to the off-line ones thanks to the large processing time available , while the latter has to rely on simplified algorithms.

2.3.2 Muon Trigger

The L1 muon trigger relies on the temporal and geometric correlation of the hits left by a muon on the different layers of RPC detectors installed in the muon spectrometer, as shown in Figure 2.18.

When a muon coming from the interaction point crosses the RPC detectors, it leaves hits on each of them: starting from the hit on the central station (also known as *pivot* plane, RPC2 in Figure 2.18) a "correlation window" (several windows are opened for several p_T thresholds) is opened on the RPC1 layer.

If a good hit (i.e. hits in both η and ϕ and in time with the hit on the pivot plane) is found on the RPC1 layer then a low-p_T muon candidate is found. The same algorithm is applied using the RPC3 plane to look for high-p_T muon candidates. Once a muon candidate is found, the RoI is propagated to the L2.

At the L2 the muon track is reconstructed for the first time: there are algorithms which reconstruct the muon tracks in the ID and in the MS separately and then combine them in order to determine of p_T, η and ϕ. At this level the p_T measurement is not done by a fit, but *look-up tables* are used: the p_T estimation is done starting from the relation

$$\frac{1}{s} = A_0 \cdot p_T + A_1 \tag{2.10}$$

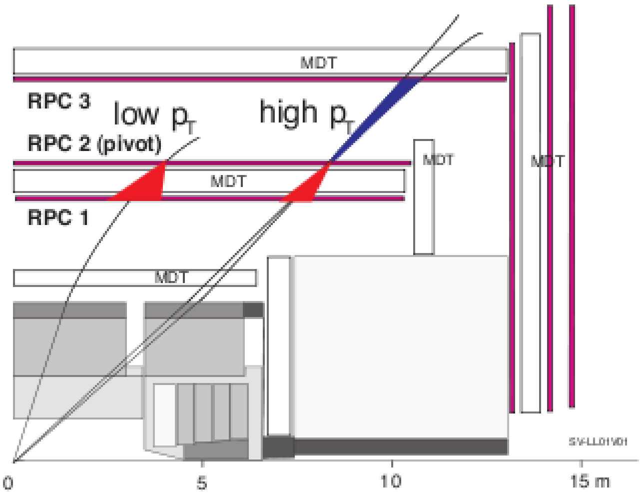
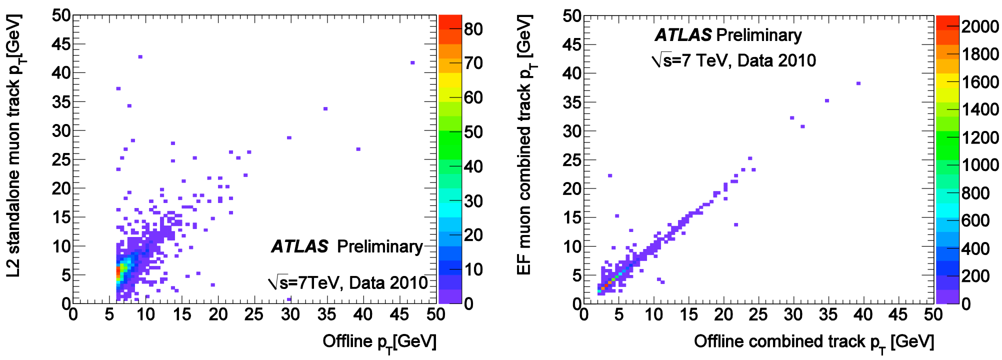


Figure 2.18: L1 muon trigger algorithm: a muon coming from the interaction point leaves hits on the three layers of RPC detectors installed in the muon spectrometer. The position of the different hits is correlated as a function of the muon p_T.

where s is the sagitta of the muon track and A_0 and A_1 are two constant values needed to take into account the magnetic field and the energy loss in the calorimeters respectively. A look-up table is basically a table whose columns and rows represent the $\eta - \phi$ segmentation of the ATLAS detector, and in each cell a (A_0, A_1) pair is contained. For each muon candidate, given η, ϕ and s, a fast estimation of the p_T is possible. This method is used since at the L2 there is not enough time to perform a real fit to precisely measure the track p_T. Once the full track is reconstructed (from the ID to the MS), the calorimetric activity around it is measured, in order to apply the isolation requirements.

At the EF the muon reconstruction algorithms perform again the operations performed by the L2 algorithms, but now the full detector with its full granularity can be accessed, and a real fit of the muon track is performed.



(a) (b)

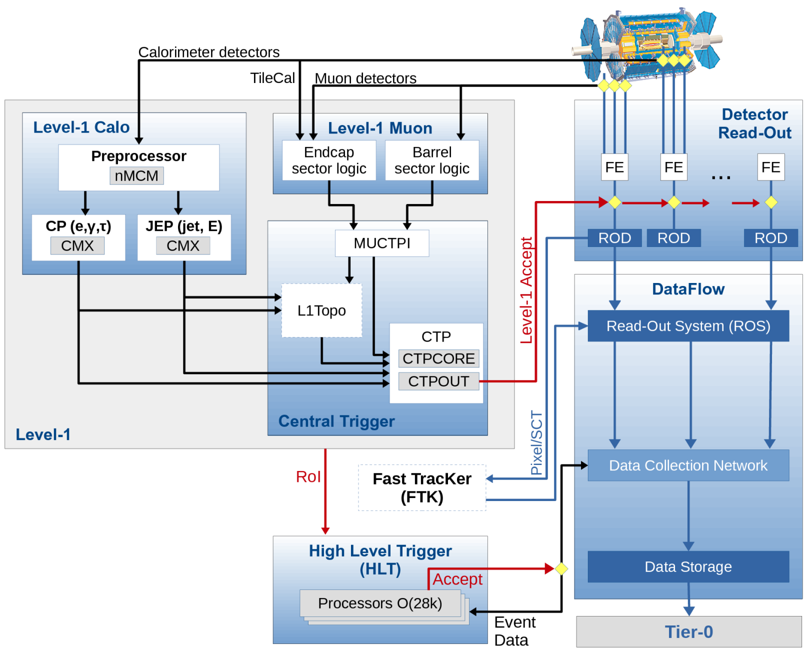
Figure 2.19: Correlation between the muon p_T reconstructed at several trigger levels (level 2 in (a) and event filter in (b)) and the off-line reconstruction.

Figure 2.19 shows the correlation between the muon p_T reconstructed at different trigger levels and the off-line reconstruction: in Figure 2.19a the correlation between the L2 stand alone p_T is shown, while in Figure 2.19b the correlation between the EF combined p_T measurement and the off-line one is shown.

As can be seen the EF measurement is much more accurate and precise compared to the one performed at L2. The corresponding plot for L1 is not shown since at L1 the muon p_T is not really measured, but, as explained above, only a threshold is available.

2.3.3 Trigger/DAQ system in Run-II

The TDAQ system used during Run 1, described in detail in [58]-[59], has been upgraded during LS1 in order to prepare for the expected higher rates in Run 2 2.20. Compared to Run 1, the LHC has increased its centre-of-mass energy from 8 to 13 TeV, and the nominal bunch-spacing has decreased from 50 to 25 ns. Due to the increase in energy, trigger rates are on average 2.0 to 2.5 times larger for the same luminosity and with the same trigger criteria. Since the bunch-spacing also decrease certain trigger rates (e.g.muons) increases due to additional interactions from neighbouring bunch-crossings (out-of-time pile-up). The main changes relevant to the trigger system are briefly described below.



(a)

Figure 2.20: The ATLAS TDAQ system in Run 2 with emphasis on the components relevant for triggering.

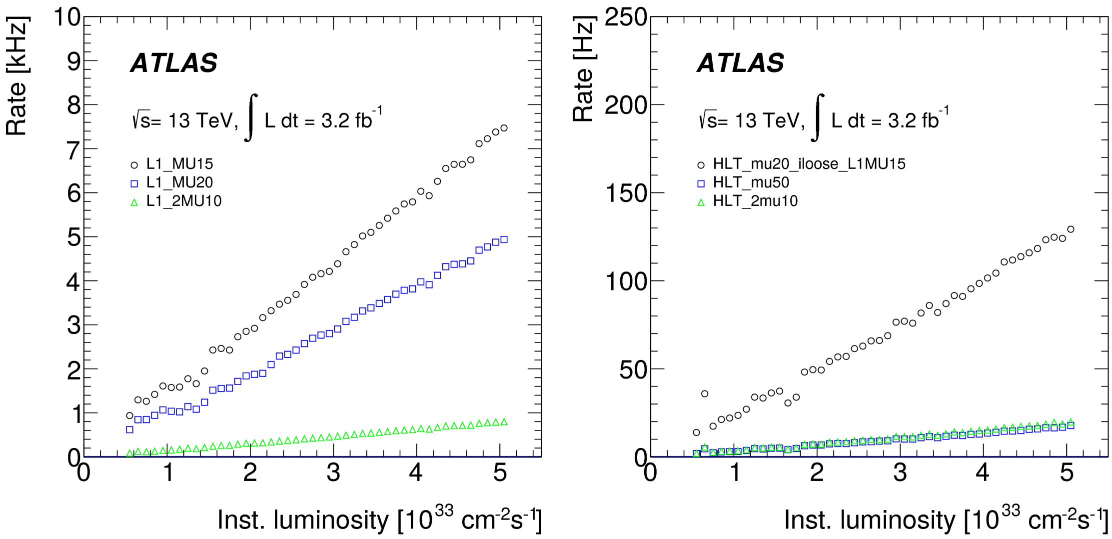
In the L1 Central Trigger, a topological trigger (L1Topo), consisting of two FPGA-based (Field-Programmable Gate Arrays) processor modules, are installed. Each modules is programmed to perform selections based on geometric or kinematic association between trigger objects received from the L1Calo or L1Muon systems. This includes the refined calculation of global event quantities such as missing transverse momentum (with magnitude EmissT). The system was fully installed and commissioned during 2016 [60]. The Muon-to-CTP interface (MUCPTI) and the CTP were upgraded to provide inputs to and receive inputs from L1Topo, respectively. In order to better address sub-detector specific requirements, the CTP now supports up to four independent complex dead-time settings operating simultaneously. In addition, the number of L1 trigger selections (512) and bunch-group selections (16), defined later, were doubled compared to Run 1.

The muon barrel trigger changed with respect to Run 1 only in the regions close to the feet that support the ATLAS detector, where the presence of support structures reduces trigger coverage. To recover trigger acceptance, a fourth layer of RPC trigger chambers was installed before Run 1 in the projective region of the acceptance holes. Only during LS1, these RPC layers were equipped with trigger electronics and they are fully operational staring from 2016. Additional chambers were installed during LS1 to cover the acceptance holes corresponding to two elevator shafts at the bottom of the muon spectrometer. The new feet and elevator chambers are expected to increase the overall barrel trigger acceptance by 2.8 and 0.8 percentage points, respectively.

2.3.4 L1 muon trigger efficiency

The lowest-threshold single-muon trigger ($mu20_iloose_L1MU15$) requires a minimum transverse momentum of 20 GeV for combined muon candidates in addition to a loose isolation: the scalar sum of the track p_T values in a cone of size $\Delta R = 0.2$ around the muon candidate is required to be smaller than 12% of the muon transverse momentum. The isolation requirement reduces the rate by a factor of approximately 2.5 with a negligible efficiency loss. The trigger is seeded by $L1_MU15$, which requires a transverse momentum above 15 GeV. At a transverse momentum above 50 GeV this trigger is complemented by a trigger not requiring isolation (mu50), to recover a small efficiency loss in the high transverse momentum region.

The lowest-threshold unprescaled di-muon trigger (2mu10) requires a minimum transverse momentum of 10 GeV for combined muon candidates. The trigger is seeded by $L1_2MU10$, which requires two muons with transverse momentum above 10 GeV. Figure 2.21 shows the rates of these triggers as a function of the instantaneous luminosity; the trigger rates scale linearly with the instantaneous luminosity.

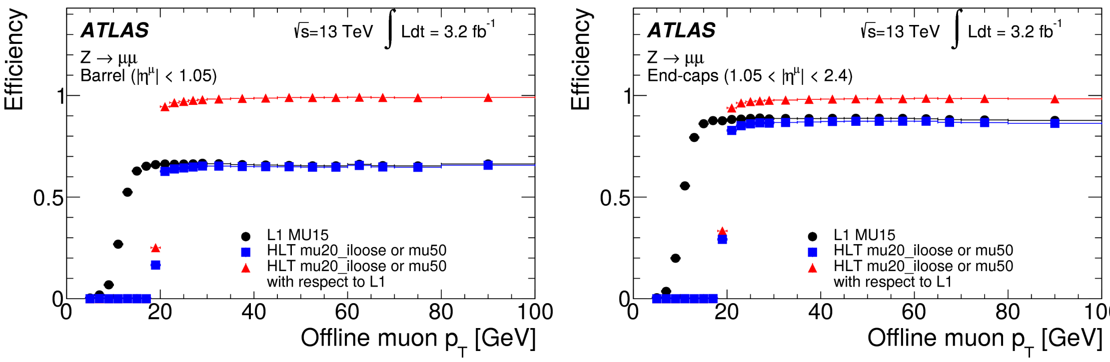


(a) (b)

Figure 2.21: (a) L1 and (b) HLT muon trigger rates as a function of the instantaneous luminosity for primary single and dimuon triggers.

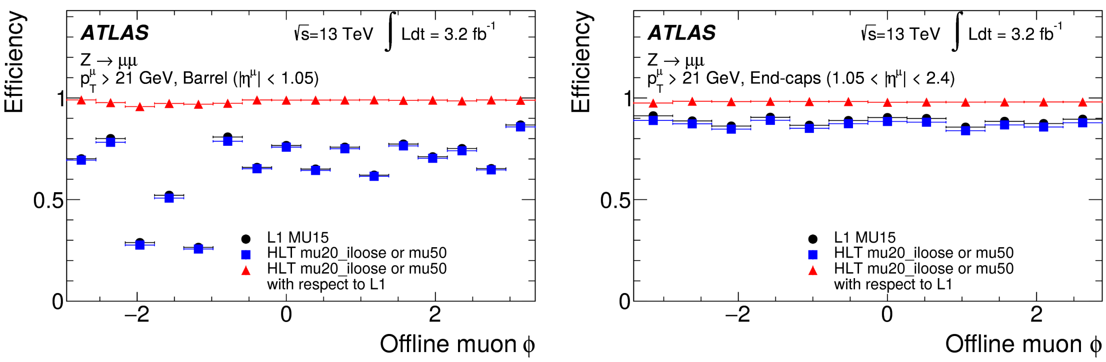
The L1 and HLT muon efficiencies are determined using a tag-and-probe method with $Z \to \mu\mu$ candidate events. Events are required to contain a pair of reference muons with opposite charge and an invariant mass within 10 GeV of the Z mass. Reference muons reconstructed offline using both ID and MS information are required to be inside the fiducial volume of the muon triggers ($|\eta| < 2.4$) and pass the medium identification requirements [61]-[62].

The absolute efficiency of the $L1_MU15$ trigger and the absolute and relative efficiencies of the logical "OR" of $mu20_ilooseandmu50$ as a function of the p_T of the offline muon track are shown in Figure 2.22. The L1 muon trigger efficiency is close to 70% in the barrel and 90% in the end-caps. The different efficiencies are due to the different geometrical acceptance of the barrel and end-cap trigger systems and local detector inefficiencies. The HLT efficiency relative to L1 is close to 100% both in the barrel and in the end-caps. Figure 2.23 shows the muon trigger efficiency as a function of the azimuthal angle ϕ of the offline muon track for (a) the barrel and (b) the end-cap regions. The reduced barrel acceptance can be seen in the eight bins corresponding to the sectors containing the toroid coils and in the two feet sectors around $\phi \sim -1.6$ and $\phi \sim -2.0$, respectively.



(a) (b)

Figure 2.22: Efficiency of the L1 muon trigger $L1_MU15$ and the HLT muon trigger $mu20_iloose_L1MU15$ OR-ed with $mu50$ as a function of the probe muon p_T, separately for (a) the barrel and (b) the end-cap regions.



(a) (b)

Figure 2.23: Efficiency of the L1 muon trigger $L1_MU15$ and the HLT muon trigger $mu20_iloose_L1MU15$ OR-ed with $mu50$ as a function of the probe muon ϕ, separately for (a) the barrel and (b) the end-cap regions.

2.3.5 RPC data format of the L1 muon trigger operation

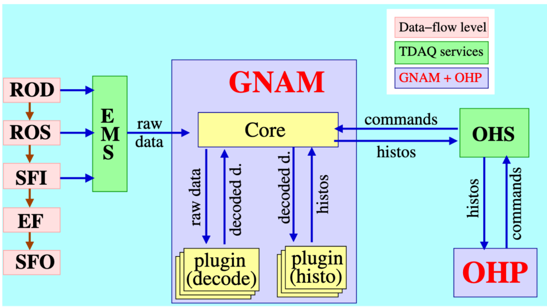
The signals from the RPC detector are amplified, discriminated and digitally shaped on the detector; indeed, at the end of the RPC strips, the Amplifier-Shaper-Discriminator (ASD) boards are attached to the chamber. In the low-p_T trigger, for each of the η and ϕ projections, the RPC signals of the first two RPC stations are sent to a Coincidence Matrix board (CM), whose core is the CM chip. This chip performs almost all of the functions needed for the trigger algorithm and also for the readout of the RPC strips. The CM board produces an output pattern containing the low-p_T results for each pair of RPC doublets in the η or ϕ projection. Moreover, the highest thresholds satisfied by the trigger conditions and a flag indicating triggers occurring in regions of overlapping RPCs are also provided. The information of two adjacent CM boards in the η projection, and the corresponding information of the two CM boards in the ϕ projection, are combined together in the low-p_T Pad Logic board. Each Pad board hosts globally 4 CMs.

In the high-p_T trigger, for each of the η and ϕ projections, the RPC signals of the third station, and the corresponding pattern result of the low-p_T trigger, are sent to a CM board, very similar to the one used in the low-p_T trigger. The information of two adjacent CM boards in the η projection, and the corresponding information of the two CM boards in the ϕ projection, are combined together in the high-p_T Pad Logic board. Again, each Pad board hosts 4 CMs. The high-p_T Pad board combines the low- and the high-p_T trigger results. The combined information is sent to a Sector Logic (SL) board, located in the USA15 counting room.

The whole RPC barrel trigger system is read-out by 32 RODs. The ROD collects the information, of both middle and outer chambers, of two adjacent half-barrel large octants. The ROD are located in the USA15 underground counting room. The format structure of the RPC data is based on the fragment provided by the CM boards. Here there are the fired RPC strips, the trigger output and the highest threshold satisfied by the trigger logic and the overlap flag. The fragments of (up to) eight CMs belonging a given Pad board (4 low-p_T CMs plus 4 low-p_T CMs) are written out one after the other, adding a header/footer at the beginning/end of this list. This is a Pad fragment. Similarly, the fragments of the Pad belonging a given Sector are written out one after the other, adding a header/footer at the beginning/end of this list. This makes the RX fragment. Finally, two adjacent RX fragments belonging to the same half-barrel trigger system are put together and form the ROD fragment.

2.3.6 Online monitoring

The TDAQ offers a number of services to implement online applications. Some of them are specially devoted to monitoring purposes, and both GNAM and OHP fully exploit them [63]; these two last service are mainly used during the online monitoring of the RPC 2.24. The Information Service (IS) is used for archiving and sharing information among the various DAQ applications; IS is typically used to route information such as run conditions, beam parameters, log messages, however IS is general enough to address any kind of information. Indeed, the Online Histograms Service (OHS) uses the IS to manage histograms, providing a transient storage between histogram producers and displays. The OHS allows also the routing of commands. Indeed, any application can issue a command related to a particular histogram, the OHS takes care of sending the command to the appropriate histogram provider. The Event Monitoring Service (EMS) provides data samplers at the ROD, ROS and SFI levels, that run in parallel with the main data taking. Monitoring applications can request events to the EMS, at a fixed rate or at the maximum possible rate; the request will be satisfied up to a rate that does not affect the data taking performance.



(a)

Figure 2.24: Schema of the GNAM and OHP monitoring chain.

GNAM is divided into two parts: the Core and the detector plugins. The Core handles the common actions, while detector specific code is implemented in the plugins. The Core parses, at run time, a configuration file, where the plugins to be loaded and the initialization parameters are specified. The Core is responsible for asking event fragments to the EMS and unpacking them, ending up with a list of ROD fragments; it cannot go further because there is no standard format for the data inside the ROD. It is up to each plugin, within its decoding routine, to find the relevant RODs and decode them. The decoded data are then collected back by the Core and provided to the histogramming plugin, where they will are analyzed and used to fill histograms. The aim of having all the data decoded before calling any histogramming routine is to allow correlation histograms. Indeed, as every histogramming function can access all the decoded data, it is easy to study the correlation of different detectors, with separate plugins, without duplicating the effort of decoding the raw data.

The ATLAS DAQ system was implemented as a Finite State Machine. A very simplified schema would comprehend these states:

- INITIAL, when the applications have just started;
- CONFIGURED, when every initialization task has been successfully completed;
- RUNNING, when the data are beingacquired.
- PAUSE, there is the possibility of pausing a run without interrupting it.

For debugging purposes GNAM has been designed to be also controllable interactively. This allows to easily start and stop the monitoring processes without interfering with the data taking.

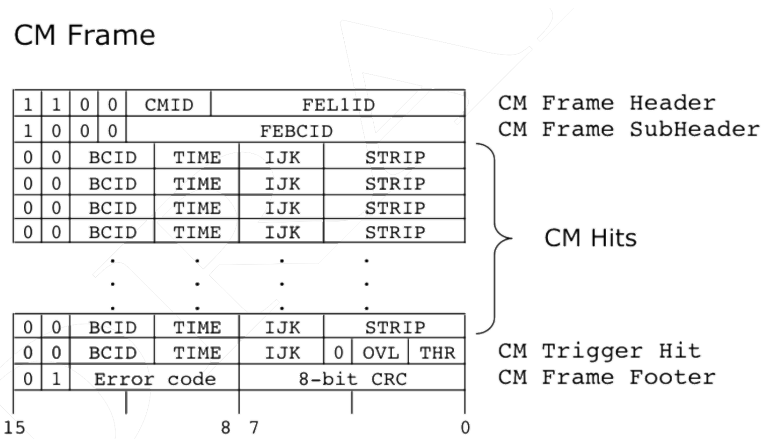
2.3.7 SEU monitoring

The detectors are in a very high radiation environment and the electronics could be affected by some issues during the activity. A kind of problem in the right electronics working is the Single Event Upset (SEU) A SEU is a change of state caused by one single ionizing particle striking a sensitive node in the electronic device.

In the RPC system we have two kind of SEU:

- it could occur on a CM or a PAD of the RPC detectors;
- it could occur in one of the many configuration register.

in the first case, the data fragment provided by the CM boards could be designed to monitor and control this problem; a flag in the footer fragment has been designed to check this case 2.25; while, in the second one, the writing of one of the headers of sub-fragments will be corrupted and there is no way to control the flow of the SEU.



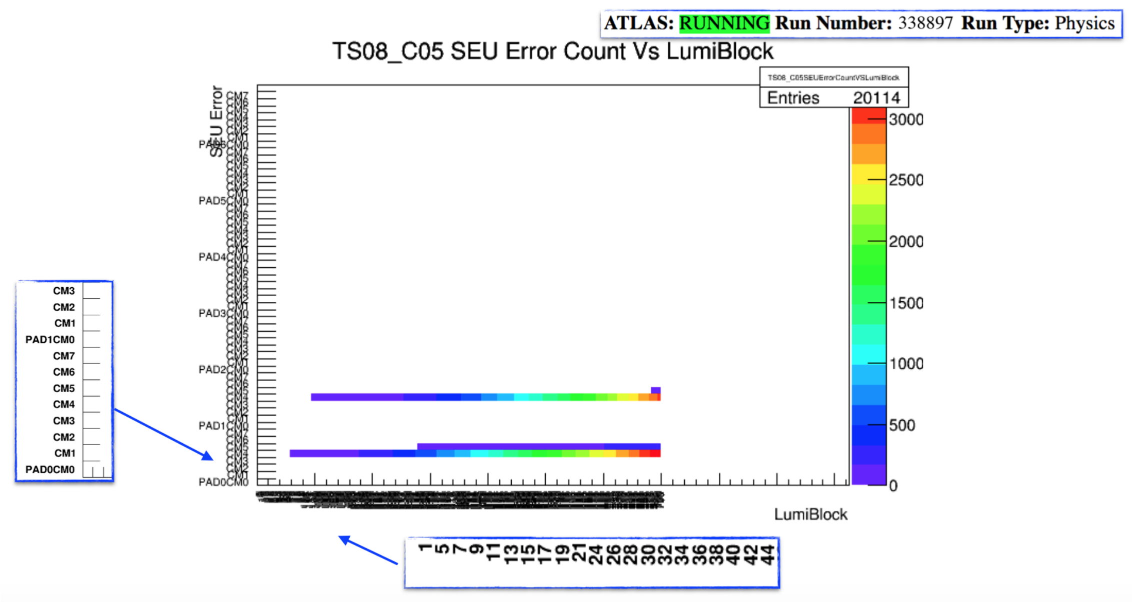
(a)

Figure 2.25: The scheme of the Coincidence Matrix read-out fragment; this is made of work of 16 bit that code all the needed information coming from the RPC.

In the framework of the GNAM application, a dedicated application for the muon barrel trigger monitoring is designed. During the 2017, few new tools for the monitoring have been added. The Figure 2.26 shows the SEU occurance in each of the CM of RPC system as a function of the LumiBlock, or in other terms of the time of the acquisition of the run. Also, the dimension of the ROD fragment distribution and the dimension of the ROD fragment as a function of the LumiBlock have been added in that period.

Bit position	Flags when set to 1
8	Almost Full latency memory
9	Almost full derandomizer
10	Almost full serializer
11	Almost full Level-1 fifo
12	Busy
13	Single event Upset (SEU)

Table 2.2: Description of the CM Error word present in frame footer.



(a)

Figure 2.26: Monitoring histograms added during the 2017 in the online routine of the L1 muon barrel trigger.

Chapter 3

Physics object reconstruction

An overview of the reconstruction and identification techniques of the main physics objects crucial in the physics analyses at a collider is provided in this chapter. The attention will focus on the objects directly used in this thesis, electrons, muons and jets. The jets identification represent an interesting challenge, indeed, their properties could be useful to the discrimination of the initial particle that initiated the jet, such as the quark/gluon identification, b-jets identification and W/Z-jets identification.

3.1 Identification and reconstruction of electrons

The performances of the reconstruction, identification and isolation algorithms are evaluated both in the data and in the MC simulation using the electrons coming from the two resonant processes $Z \to ee$ and $J/\Psi \to ee$ [64].

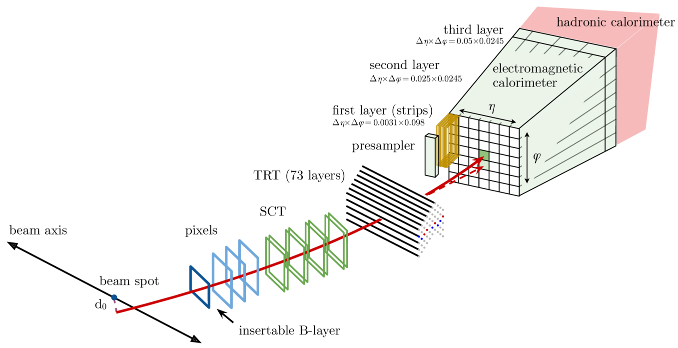
Electrons produced isolated from other SM particles in the pp collisions are important pieces in analyses of SM measurements or searches of new physics BSM. However, the experimental spectra of the electrons must be corrected for the selection efficiencies, such as those related to the trigger, as well as particle isolation, identification, and reconstruction, before absolute measurements can be made. These efficiencies are estimated directly from data using tag-and-probe methods. According to this method, unbiased samples of electrons (probes) are selected by using strict selection requirements on the second object (tags) produced from the particle's decay of known resonances such as Z or J/Ψ decaying into a pair of electrons. The events are selected on the basis of the electron-positron invariant mass. The efficiency of a given requirement can then be determined by applying it to the probe sample after accounting for residual background contamination.

The total efficiency ϵ_{tot} can be factorised as a product of different efficiency terms:

$$\epsilon_{tot} = \epsilon_{EMclus} \cdot \epsilon_{reco} \cdot \epsilon_{id} \cdot \epsilon_{iso} \cdot \epsilon_{trig} = \left(\frac{N_{cluster}}{N_{all}}\right) \cdot \left(\frac{N_{reco}}{N_{cluster}}\right) \cdot \left(\frac{N_{id}}{N_{reco}}\right) \cdot \left(\frac{N_{iso}}{N_{id}}\right) \cdot \left(\frac{N_{trig}}{N_{iso}}\right) \quad (3.1)$$

where N_{all} is the number of produced electrons, $N_{cluster}$ is the number of reconstructed EM cluster, N_{reco} is the number of reconstructed electrons candidates, N_{id} is the number of identified and reconstructed electrons candidates, N_{iso} is the number of electrons candidates applying isolation, identification and reconstructed requirements, N_{trig} is the number of triggered electron candidates (after applying isolation, identification and reconstruction requirements). The tag-and-probe method is applied in the same way for data and simulated events.

An electron coming from the interaction point traverses the detector material and loses a significant amount of its energy due to bremsstrahlung. The radiated photon may convert into an electron-positron pair which itself can interact with the detector material. These positrons, electrons, and photons are usually emitted in a very collimated fashion and are normally reconstructed as part of the same electromagnetic cluster. These interactions can occur inside the inner-detector volume or even in the beam pipe, generating multiple tracks in the inner detector, or can instead occur downstream of the inner detector, only impacting the shower in the calorimeter. As a result, it is possible to produce and match multiple tracks to the same electromagnetic cluster, all originating from the same primary electron.



(a)

Figure 3.1: A schematic illustration of the path of an electron through the detector. The red trajectory shows the hypothetical path of an electron, which first traverses the tracking system (pixel detectors, then silicon-strip detectors and lastly the TRT) and then enters the electromagnetic calorimeter. The dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system.

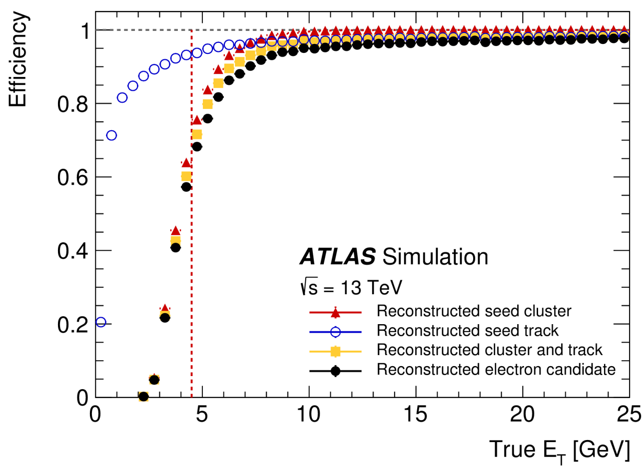
The reconstruction of electron candidates uses the high-granularity electromagnetic calorimeter and the inner detector and it is based on three fundamental components characterising the signature of electrons: localised clusters of energy deposits in the electromagnetic calorimeter, charged-particle tracks identified in the inner detector and close matching of the tracks to the clusters in the $\eta - \phi$ space to form the final electron candidates. Figure 3.1 shows a schematic picture of the elements that enter into the reconstruction and identification of an electron. Therefore, electron reconstruction in the precision region of the ATLAS detector ($|\eta| < 2.47$) proceeds along the previous three steps:

- *Seed-cluster reconstruction*: Electromagnetic-energy cluster candidates are seeded from localised energy deposits using a sliding-window algorithm [65] of size 3×5 towers (each element of the 200×256 grid in which is divided the calorimeter) in the $\eta\times\phi$. The centre of the 3×5 seed cluster moves in steps of 0.025 in both two dimension, searching for localised energy deposits.

- *Track reconstruction*: Track seeds are formed from sets of three space-points in the silicon-detector layers. The track reconstruction then proceeds in three steps: pattern recognition, ambiguity resolution and TRT extension [66]. Track candidates with $p_T > 400\ MeV$ are fit, according to the hypothesis used in the pattern recognition, using the ATLAS Global χ^2 Track Fitter [67]. A subsequent fitting procedure, using an optimised Gaussian-sum filter (GSF) [68], designed to better account for energy loss of charged particles in material, is applied to the clusters of raw measurements.

- *Electron-candidate reconstruction* A matching of the calo-cluster and of the track with $-0.10 < \Delta\phi(track, cluster) < 0.05$ is done; further algorithms are used to improve the ambiguity in case of multiple matching and of electron conversion [64]. Finally, reconstructed clusters are formed around the seed clusters using an extended window of size 3×7 in the barrel region or 5×5 in the endcap. The energy of the clusters must ultimately be calibrated to correspond to the original electron energy. This detailed calibration is performed using multivariate techniques [69]-[70].

In Figure 3.2 the total reconstruction efficiency for simulated electrons as a function of the true (generator) transverse energy E_T for each step of the electron-candidate formation is showed.



(a)

Figure 3.2: Total reconstruction efficiency for simulated electrons as a function of the true (generator) transverse energy E_T for each step of the electron-candidate formation: seed-cluster reconstruction (red triangles), seed-track reconstruction using the Global χ^2 Track Fitter (blue open circles), both of these steps together but instead using GSF tracking (yellow squares), and the final reconstructed electron candidate, which includes the track-to-cluster matching (black closed circles). The total reconstruction efficiency is less than 60% below 4.5 GeV (dashed line).

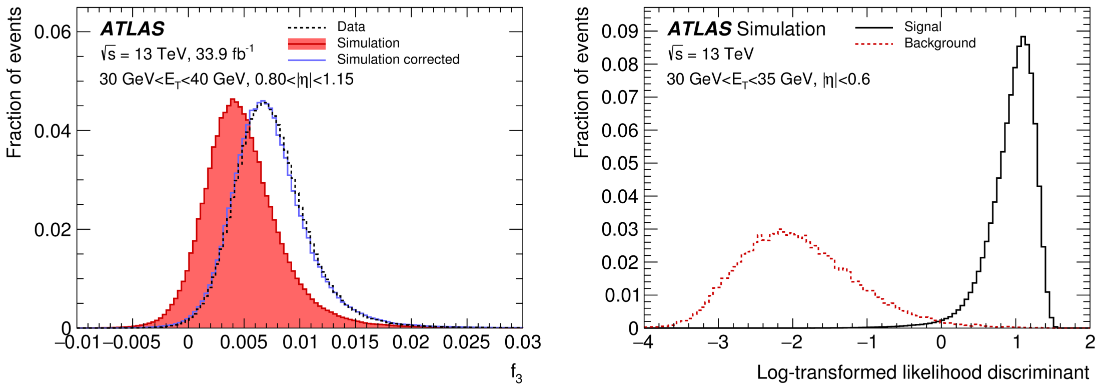
Prompt electrons in the central region of the detector ($|\eta| < 2.47$) are selected using a likelihood-based (LH) identification. The inputs to the LH include informations from the tracking system or the calorimeter system and quantities that combine both of them (all inputs is reported here [64]). The method is based on a pseudo-likelihood function that combines all the p.d.f. of the inputs, neglecting all the correlations:

$$L_{S(B)}(\bar{x}) = \prod_{i=1}^{N} pdf_{S(B)}(x_i) \tag{3.2}$$

the signal is prompt electrons while the background is the combination of jets that are similar to the signature of prompt electrons, electrons from photon conversions in the detector material and non-prompt electrons from the decay of hadrons containing heavy flavours. The final discriminant used (Figure 3.3-b) is given by an inverse sigmoid of the likelihood ratio:

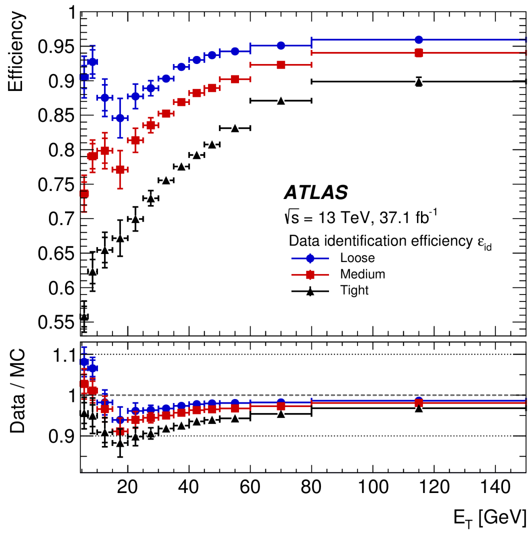
$$d'_L = -\tau^{-1} ln(d_L^{-1} - 1) \qquad with \qquad d_L = \frac{L_S}{L_S + L_B} \tag{3.3}$$

To cover the various required prompt-electron signal efficiencies and corresponding background rejection factors needed by the physics analyses carried out within the ATLAS Collaboration, four fixed values of the LH discriminant are used to define four operating points. These operating points are referred to as *VeryLoose*, *Loose*, *Medium* and *Tight* and correspond to increasing thresholds for the LH discriminant. As an example, the efficiencies for identifying a prompt electron with $E_T = 40$ GeV are 93%, 88%, and 80% for the Loose, Medium, and Tight operating points, respectively. The efficiencies curves as a function of the energy of the electrons for the different operating points are showed in Figure 3.4.



(a) (b)

Figure 3.3: The f_3 distribution for data and simulated events obtained using the $Z \to ee$ tag-and-probe method. The simulation is shown before (shaded histogram) and after (open histogram) applying a correction (constant shift and a width-scaling factor are applied for the several input variables). The transformed LH-based identification discriminant d'_L for reconstructed electron candidates for the signal (red) and for the background (black) (right).

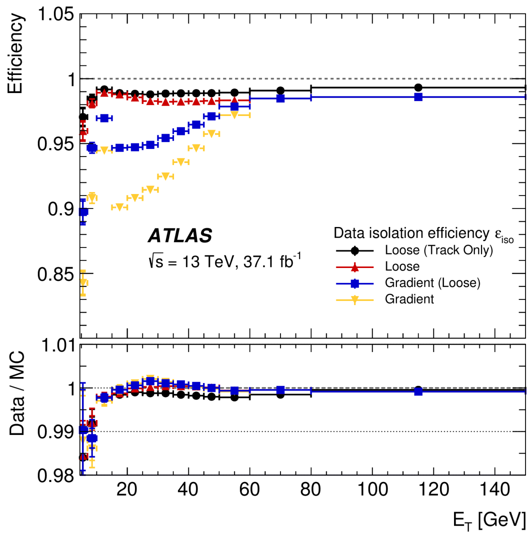


(a)

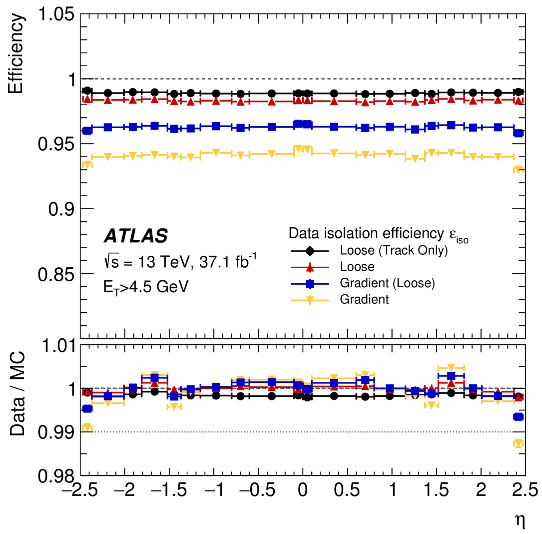
Figure 3.4: Measured LH electron-identification efficiencies in $Z \to ee$ events for the Loose (blue circle), Medium (red square) and Tight (black triangle) operating points as a function of E_T . The bottom panels shows the data-to-simulation ratios.

The final requirement on the isolation of the reconstructed electron could be crucial in analysis that would not cover their signal electrons with similar signatures such as those coming from the production of boosted particles decaying, for example, into collimated electron-positron pairs or the production of prompt electrons, muons, and photons within a busy experimental environment such as in $t\bar{t}$ production can obscure the picture.

Variables are constructed that quantify the amount of activity in the near space of the candidate object, something usually performed by summing the transverse energies of clusters in the calorimeter (E_T^{isol}) or the transverse momenta of tracks (p_T^{isol}) in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi^2)}$ around the direction of the electron candidate, excluding the candidate itself. Figure 3.5 shows the isolation efficiencies measured in data and the corresponding data-to-simulation ratios as a function of the electron E_T and η for the operating points given in Table 3.6 and for candidate electrons satisfying Tight identification requirements. The efficiencies that determine the values of the requirements given in Table 3.6 are evaluated in simulation from a $J/\Psi \to ee$ sample for $E_T < 15\ GeV$ and from a $Z \to ee$ for $E_T > 15\ GeV$.



(a)



(b)

Figure 3.5: Isolation efficiencies for data (in the upper panels) and the ratio to simulation (lower panels) for the operating points given in Table 3.6 as a function of candidate-electron E_T (left) and η (right).

Operating point	$E_{\mathrm{T,cone}}^{\mathrm{isol}}$ ($\Delta R = 0.2$)	$p_{\mathrm{T,var}}^{\mathrm{isol}}$ ($R_{\mathrm{max}} = 0.2$)	Total ϵ_{iso}
Loose (Track Only)	-	$\epsilon_{\mathrm{iso}} = 99\%$	99%
Loose	$\epsilon_{\mathrm{iso}} = 99\%$	$\epsilon_{\mathrm{iso}} = 99\%$	98%
Gradient	$\epsilon_{\mathrm{iso}} = 0.1143 \times p_{\mathrm{T}} + 92.14\%$	$\epsilon_{\mathrm{iso}} = 0.1143 \times p_{\mathrm{T}} + 92.14\%$	90(99)% at 25(60) GeV
Gradient (Loose)	$\epsilon_{\mathrm{iso}} = 0.057 \times p_{\mathrm{T}} + 95.57\%$	$\epsilon_{\mathrm{iso}} = 0.057 \times p_{\mathrm{T}} + 95.57\%$	95(99)% at 25(60) GeV
Fix (Loose)	$E_{\mathrm{T,cone}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.20$	$p_{\mathrm{T,var}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.15$	-
Fix (Tight)	$E_{\mathrm{T,cone}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.06$	$p_{\mathrm{T,var}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.06$	-
Fix (Tight, Track Only)	-	$p_{\mathrm{T,var}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.06$	-
Fix (Calo Only)	$E_{\mathrm{T,cone}}^{\mathrm{isol}} < 3.5$ GeV	-	-
Fix (Track $R_{\mathrm{max}} = 0.4$)	$E_{\mathrm{T,cone}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.11$	$p_{\mathrm{T,var}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.06$	-

(a)

Figure 3.6: Definition of the electron-isolation operating points and isolation efficiency ϵ_{iso}. For the "Gradient" operating points, the units of p_T are GeV. All operating points use a cone size of $\Delta R = 0.2$ for calorimeter isolation and $R_{max} = 0.2$ for track isolation except for the final entry "Fix (Track)" which uses $R_{max} = 0.4$.

3.2 Identification and reconstruction of muons

During the 2013-2015 shutdown of the LHC, the ATLAS detector was equipped with additional muon chambers and a new innermost Pixel layer, the *Insertable B-Layer*, providing measurements closer to the interaction point. Moreover, the muon reconstruction software was updated and improved. The results in the following are based on the analysis of a large sample of $J/\Psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays reconstructed in 3.2 fb^{-1} of pp collisions recorded in 2015 [62].

Muon reconstruction is first performed independently in the ID and MS. The information from individual sub-detectors is then combined to form the muon tracks that are used in physics analyses. Muon reconstruction in the MS starts with a search for hit patterns inside each muon chamber to form segments. In each MDT chamber and nearby trigger chamber, a Hough transform is used to search for hits aligned on a trajectory in the bending plane of the detector. The MDT segments are reconstructed by performing a straight-line fit to the hits found in each layer. The RPC or TGC hits measure the coordinate orthogonal to the bending plane. Segments in the CSC detectors are built using a separate combinatorial search in the η and ϕ detector planes. Muon track candidates are then built by fitting together hits from segments in different layers. The algorithm used for this task performs a segment-seeded combinatorial search that starts by using as seeds the segments generated in the middle layers of the detector where more trigger hits are available. The search is then extended to use the segments from the outer and inner layers as seeds. The segments are selected using criteria based on hit multiplicity and fit quality and are matched using their relative positions and angles. The hits associated with each track candidate are fitted using a global χ^2 fit. A track candidate is accepted if the χ^2 of the fit satisfies the selection criteria.

The combined ID–MS muon reconstruction is performed according to various algorithms based on the information provided by the ID, MS, and calorimeters. Four muon types are defined depending on which sub-detectors are used in the reconstruction:

- *Combined (CB) muon*: track reconstruction is performed independently in the ID and MS, and a combined track is formed with a global refit that uses the hits from both the ID and MS sub-detectors. Most muons are reconstructed following an outside-in pattern recognition, in which the muons are first reconstructed in the MS and then extrapolated inward and matched to an ID track. An inside-out combined reconstruction, in which ID tracks are extrapolated outward and matched to MS tracks, is used as a complementary approach.
- *Segment-tagged (ST) muons*: a track in the ID is classified as a muon if, once extrapolated to the MS, it is associated with at least one local track segment in the MDT or CSC chambers. ST muons are used when muons cross only one layer of MS chambers, either because of their low p_T or because they fall in regions with reduced MS acceptance.
- *Calorimeter-tagged (CT) muons*: a track in the ID is identified as a muon if it can be matched to an energy deposit in the calorimeter compatible with a minimum-ionizing particle. This type has the lowest purity of all the muon types but it recovers acceptance in the region where the ATLAS muon spectrometer is only partially instrumented to allow for cabling and services to the calorimeters and inner detector. The identification criteria for CT muons are optimized for that region ($|\eta| < 0.1$ and a momentum range of $15 < p_T < 100\, GeV$.
- *Extrapolated (ME) muons*: the muon trajectory is reconstructed based only on the MS track and a loose requirement on compatibility with originating from the IP. The parameters of the muon track are defined at the interaction point, taking into account the estimated energy loss of the muon in the calorimeters. In general, the muon is required to traverse at least two layers of MS chambers to provide a track measurement, but three layers are required in the forward region. ME muons are mainly used to extend the acceptance for muon reconstruction into the region $2.5 < |\eta| < 2, 7$, which is not covered by the ID.

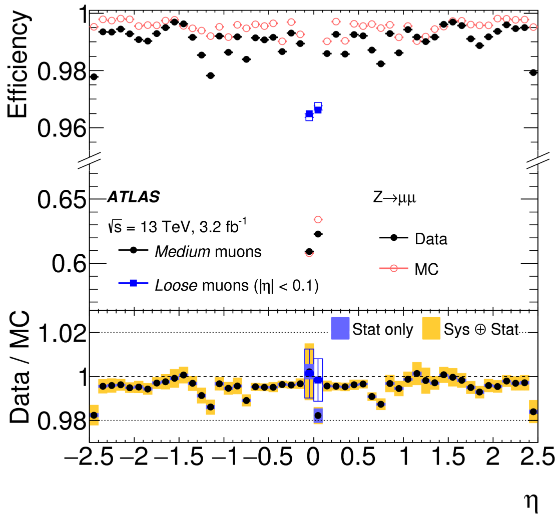
Muon identification is performed by applying quality requirements that suppress background, mainly from pion and kaon decays, while selecting prompt muons with high efficiency and/or guaranteeing a robust momentum measurement. Several variables offering good discrimination between prompt muons and background muon candidates are studied. For CB tracks, the variables used in muon identification are:

- $q/p\, significance$, defined as the absolute value of the difference between the ratio of the charge and momentum of the muons measured in the ID and MS divided by the sum in quadrature of the corresponding uncertainties;
- ρ defined as the absolute value of the difference between the transverse momentum measurements in the ID and MS divided by the p_T of the combined track;
- normalized χ^2 of the combined track fit.

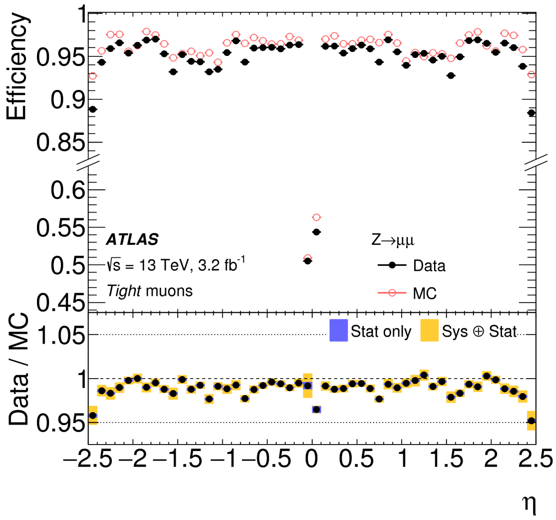
Four muon identification selections (*Medium*, *Loose*, *Tight* and *High-p_T*) are provided to address the specific needs of different physics analyses. Loose, Medium and Tight are inclusive categories in that muons identified with tighter requirements are also included in the looser categories.
The different selections are:

- **Medium muons**: the Medium identification criteria provide the default selection for muons in ATLAS. This selection minimizes the systematic uncertainties associated with muon reconstruction and calibration. Only CB and ME tracks are used. The former are required to have ≥ 3 hits in at least two MDT layers, except for tracks in the $|\eta| < 0,1$ region, where tracks with at least one MDT layer but no more than one MDT hole layer are allowed. The latter are required to have at least three MDT/CSC layers, and are employed only in the $2.5 < |\eta| < 2,7$ region to extend the acceptance outside the ID geometrical coverage. A loose selection on the compatibility between ID and MS momentum measurements is applied to suppress the contamination due to hadrons misidentified as muons.
- **Loose muons**: the Loose identification criteria are designed to maximise the reconstruction efficiency while providing good-quality muon tracks. They are specifically optimized for reconstructing Higgs boson candidates in the four-lepton final state . All muon types are used. All CB and ME muons satisfying the Medium requirements are included in the Loose selection. CT and ST muons are restricted to the $|\eta| < 0,1$ region. In the region $|\eta| < 2,5$ about 97.5% of the Loose muons are combined muons, approximately 1.5% are CT and the remaining 1% are reconstructed as ST muons.
- **Tight muons**: tight muons are selected to maximize the purity of muons at the cost of some efficiency. Only CB muons with hits in at least two stations of the MS and satisfying the Medium selection criteria are considered. The normalized χ^2 of the combined track fit is required to be < 8 to remove pathological tracks. A two-dimensional cut in the ρ and q/p significance variables is performed as a function of the muon p_T to ensure stronger background rejection for momenta below $20\, GeV$ where the misidentification probability is higher.
- **High-p_T muons**: the High-p_T selection aims to maximize the momentum resolution for tracks with transverse momentum above $100\, GeV$. The selection is optimized for searches for high-mass Z' and W' resonances. CB muons passing the Medium selection and having at least three hits in three MS stations are selected. Specific regions of the MS where the alignment is sub-optimal are vetoed as a precaution. Requiring three MS stations, while reducing the reconstruction efficiency by about 20%, improves the p_T resolution of muons above $1.5\, TeV$ by approximately 30%.

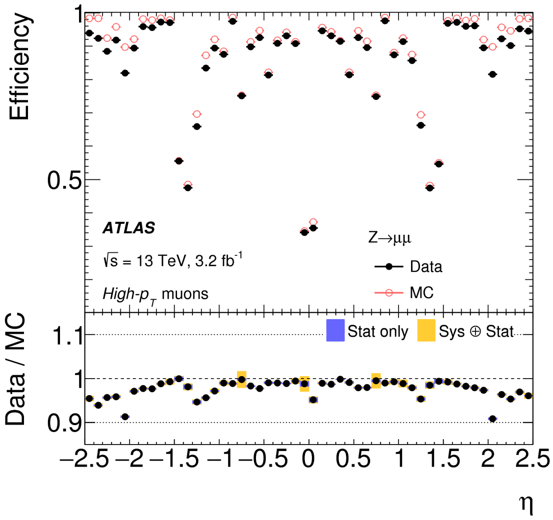
The muon reconstruction efficiency as a function of η is showed in Figure 3.7 for the four reconstruction criteria.



(a)



(b)



(c)

Figure 3.7: Muon reconstruction efficiency as a function of η measured in $Z \to \mu\mu$ events for muons with $p_T > 10$ GeV shown for the Medium muon selection (a), Tight (b) and High-p_T (c). In addition, the plot (a) also shows the efficiency of the Loose selection (squares) in the region $|\eta| < 0.1$ where the Loose and Medium selections differ significantly. The error bars on the efficiencies indicate the statistical uncertainty. Panels at the bottom show the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.

Muons originating from the decay of heavy particles, such as W,Z, or Higgs bosons, are often produced isolated from other particles, similar to the electrons as discussed in the previous section. The measurement of the detector activity around a muon candidate, referred to as muon isolation, is therefore a powerful tool for background rejection in many physics analyses.

Two variables are defined to parametrise the muon isolation, using the track or the calorimetric information. The track-based isolation variable, $p_T^{varcone30}$, is defined as the scalar sum of the transverse momenta of the tracks with $p_T > 1\ GeV$ in a cone of size $\Delta R = min(10\ GeV \cdot p_T^{\mu}, 0.3)$ around the muon of transverse momentum p_T^{μ}, excluding the muon track itself. The calorimeter-based isolation variable, $E_T^{varcone30}$, is defined as the sum of the transverse energy of topological clusters [71] in a cone of size $\Delta R = 0.2$ around the muon, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects. Seven isolation selection criteria (isolation working points) are defined, each optimized for different physics analyses (table 3.8).

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

(a)

Figure 3.8: Definition of the seven isolation working points. The discriminating variables are listed in the second column and the criteria used in the definition are reported in the third column.

Figure 3.9 shows the isolation efficiency measured for Medium muons in data and simulation as a function of the muon p_T for the LooseTrackOnly, Loose, GradientLoose and FixedCutLoose working points, with the respective data/MC ratios included in the bottom panel. The systematic uncertainties in the SFs are estimated by varying the background contributions within their uncertainties and by varying some of the selection criteria, such as the invariant mass selection window, the isolation of the tag muon, the minimum quality of the probe muon, the opening angle between the two muons, and the ΔR between the probe muon and the closest jet.

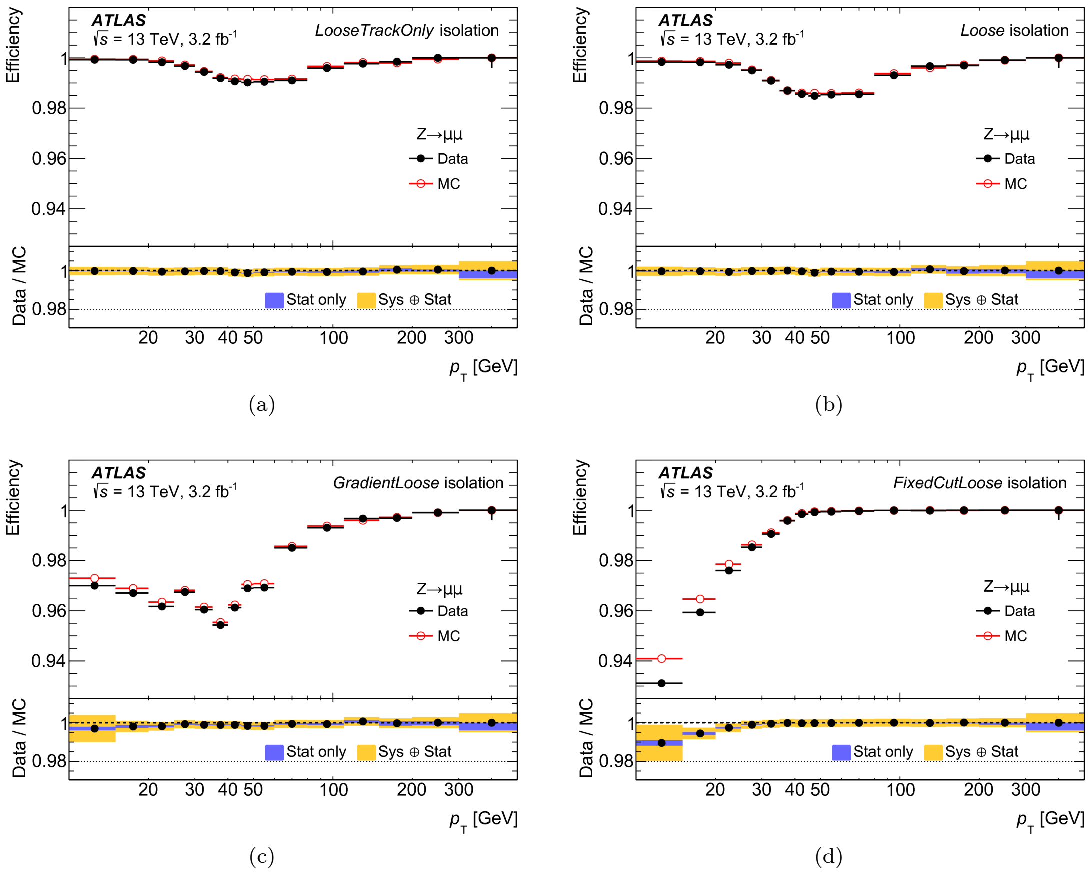


Figure 3.9: Isolation efficiency for the *LooseTrackOnly* (a), *Loose* (b), *GradientLoose* (c) and *FixedCutLoose* (d) muon isolation working points as a function of the muon transverse momentum p_T. The full (empty) markers indicate the efficiency measured in data (MC) samples. The errors shown on the efficiency are statistical only. The bottom panel shows the ratio of the efficiency measured in data and simulation, as well as the statistical uncertainties and combination of statistical and systematic uncertainties.

3.3 Jets reconstruction

At high energy pp collisions the presence of partons is overwhelming. Due to colour confinement the partons hadronize. While the resulting bunch of particles passes through the ATLAS detector, they produce tracks in the ID and energy deposits inside the calorimeters. These detector signals allow the reconstruction of *track jets* (reconstructed using track information) and *calorimeter jets* (reconstructed using calorimeter information).

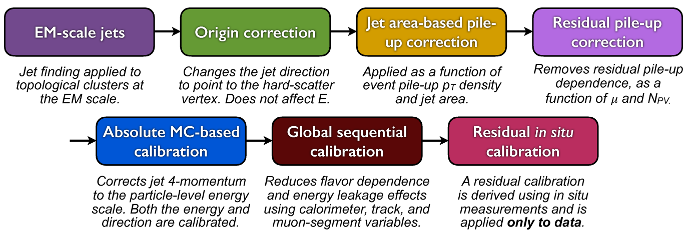
Further, the dimension of the cone of jets has been crucial in this work. Default value of $R = 0.4$ is used in jets definition for analyses in ATLAS but higher dimension, $R = 1.0$, became important when the aim of the jets is to reconstruct the boosted products decay of a weak boson, W, Z, or the higgs boson or the decay of a top quark [72]. We talk of the first type as small-R jets and as large-R jets for the second type.

This section will focus only on the jets used in this work; the small-R jets are based only on calorimeter information (EMTopo) while the large-R jets on combination of both tracks and calorimeters information (TrackCaloClusters, TCC). The reconstruction process consists in three steps: the definition of calorimeter/tracks signals, the use of a jet reconstruction algorithm to group the signals and finally the jet calibration which corrects the jet energy and momentum for the effects of ATLAS calorimeters non-compensation, dead material, leakage, out of cone and other thresholds effects.

3.3.1 Small-R jets

The calorimeter jets used in the following studies are reconstructed at the electromagnetic energy scale (EM scale) with the anti-kt algorithm [73] and radius parameter $R = 0.4$ using the FastJet 2.4.3 software package [74]. A collection of three-dimensional, massless, positive-energy topological clusters (topo-clusters) [75]-[71] made of calorimeter cell energies are used as input to the anti-kt algorithm. Topo-clusters are built from neighboring calorimeter cells containing a significant energy above a noise thresh-old that is estimated from measurements of calorimeter electronic noise and simulated pile-up noise. The calorimeter cell energies are measured at the EM scale, corresponding to the energy deposited by electromagnetically interacting particles. Jets are reconstructed with the anti-kt algorithm if they pass a p_T threshold of 7 GeV.

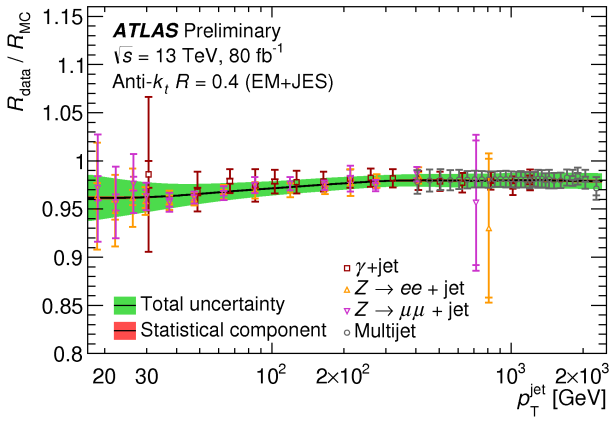
An overview of the ATLAS calibration scheme for EM-scale calorimeter jets is showed in Figure 3.10. This calibration restores the jet energy scale to that of truth jets reconstructed at the particle-level energy scale.



(a)

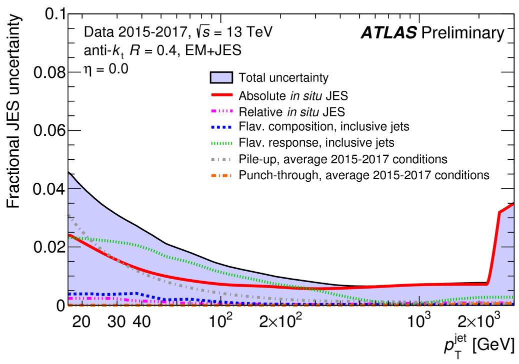
Figure 3.10: Calibration stages for EM-scale jets. Other than the origin correction, each stage of the calibration is applied to the four-momentum of the jet.

The absolute Jet Energy Scale (JES) calibration corrects the jet four-momentum to the particle-level energy scale, as derived using truth jets in dijet MC events. Further improvements to the reconstructed energy and related uncertainties are achieved through the use of calorimeter, MS, and track-based variables in the global sequential calibration. Finally, a residual in situ calibration is applied to correct jets in data using well-measured reference objects, including photons, Z bosons, and calibrated jets.

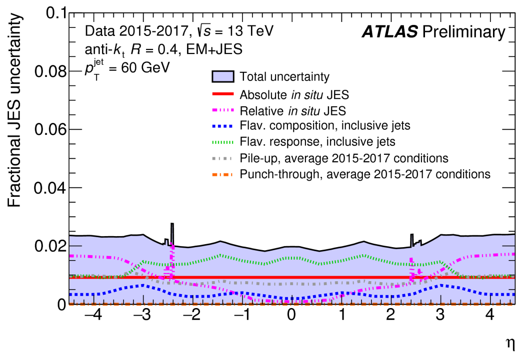


(a)

Figure 3.11: Data-to-simulation ratio of the average jet p_T response as a function of jet p_T. The combined result is based on three in situ techniques: the Z+jet balance method (dielectron channel, upward-pointing triangles; and dimuon channel, downward-pointing triangles), $\gamma + jet$ balance method (open squares) and the multijet balance (open circles). The errors represent the statistical (inner error bars and small inner band) and the total uncertainty (statistical and systematic uncertainties added in quadrature, outer error bars and outer band).



(a)



(b)

Figure 3.12: Fractional jet energy scale systematic uncertainty components as a function of jet p_T for R=0.4 $anti-k_t$ jets at $\eta = 0.0$ (a) and as a function of η for R=0.4 $anti-k_t$ jets with $p_T = 60$ GeV, reconstructed from electromagnetic-scale topo-clusters. The total uncertainty (all components summed in quadrature) is shown as a filled region topped by a solid black line. Topology-dependent components are shown under the assumption of a dijet flavour composition.

3.3.2 Large-R Jets

High energy pp collision can result in the production of massive particles, as boson W, Z, H or as the top quark, with high p_T. When these particles decay in 2 or 3 body, the typical angular separation between their decay products scales as inversely proportional to the p_T of the initial particle:

$$\Delta R \sim \frac{2m}{p_T} \tag{3.4}$$

when the decay channel is the hadronic one and when the momentum is high enough, all the products set could be reconstructed as a single large-radius (large-R) jet. These jets may have a distinct radiation pattern respect to light-quark- or gluon-initiated jets with the kinematics. In particular, the 2-body or 3-body decay of hadronically decaying W, Z, H bosons and top quarks result in a characteristic multi-prong jet substructure.

Different algorithms are used in ATLAS for the large-R jets reconstruction [76].

- *Topo clusters based (LCTopo)*, have three-dimensional clusters of calorimeter cells designed to suppress fluctuations due to noise and pileup. Topoclusters used to reconstruct the large-R jets include an additional calibration using the Local Cell Weighting (LCW) scheme, the clusters are calibrated to the hadronic scale to take into accounts for the non-compensation of the calorimeter, out-of-cluster energy, and energy falling in the dead material within the detector. Finally, the angular coordinates (η and ϕ) of the cluster are corrected for the selected primary vertex.

- *TrackCaloClusters (TCC)* implements an algorithm that also starts by matching all reconstructed ID tracks to topoclusters in the calorimeter. In the case where a single track matches a single topocluster, the p_T of the TCC object is taken from the topocluster, while the angular information is taken from the track. In more complex cases where multiple tracks are matched to multiple topoclusters, multiple TCC objects are created where each TCC object is given some fraction of the momentum of the topocluster(s). In those cases, the fraction of momenta is determined from the momenta of the matched tracks, and the angular information is still taken from the individual tracks. Like PFlow, unmatched topoclusters are included in the TCC objects unmodified, and charged TCCs (TCCs which have an associated track) which are not matched to the primary vertex are removed.

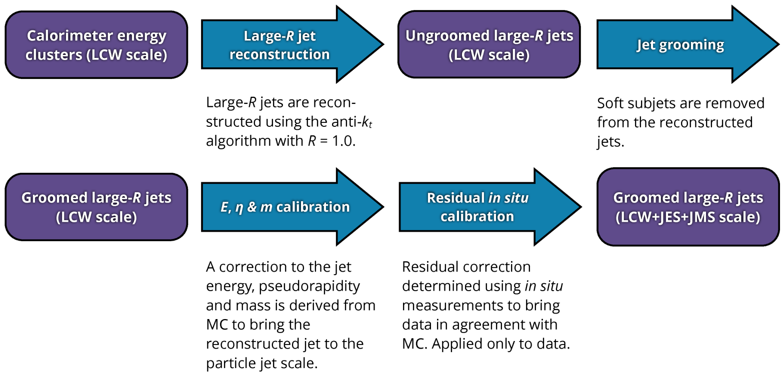
Other algorithms are developed from the ATLAS collaboration, such as Particle-Flow based algorithms [76], but they are not directly used in this work.

The different stages of the large-R jet calibration procedure are illustrated in Figure 3.13. The trimmed large-R jets are calibrated to the energy scale of stable final-state particles using corrections based on simulations. This jet-level correction is referred to as the simulation-based calibration and includes a correction to the jet mass [77]. Finally, the jets are calibrated in situ using response measurements in data.

The calibration is done taken into account the response of the detector in the measure of the energy (JES) and of the mass (JMS) of the jets (Figures 3.14); the procedure on the JES is simialr to what is done on the small-R jets. Uncertainties in the JES and JMS are derived by propagating uncertainties from the individual in situ response measurements through the statistical combination of the response in data and MC (Figure 3.15).

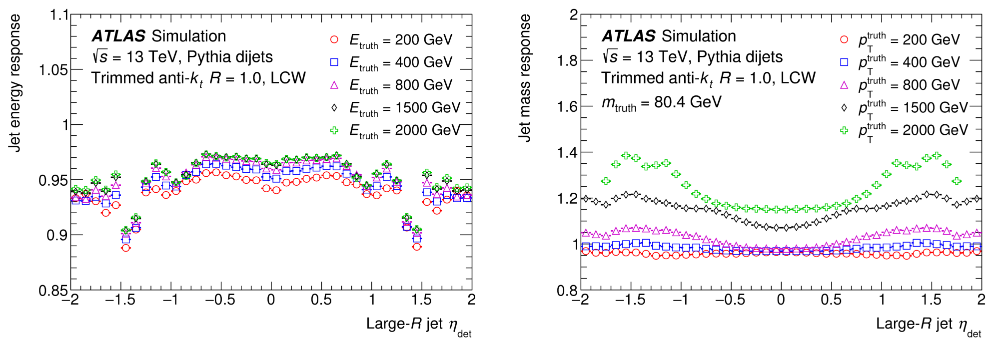
The jet mass resolution is improved by combining the jet mass measurement in the calorimeter with the measurement of the charged component of the jet within the ID [78]. A track jet is reconstructed from ID tracks with $p_T > 500\ MeV$ which are ghost-associated [79] to the topo-cluster large-Rjet. The measurement of this track jet's mass is multiplied by the ratio of the transverse momenta of the calorimeter jet and the track jet to obtain the track-assisted mass (m^{TA}):

$$m^{TA} = m^{track} \frac{p_T^{calo}}{p_T^{track}} \tag{3.5}$$



(a)

Figure 3.13: Overview of the large-R jet reconstruction and calibration procedure described in this paper. The calorimeter energy clusters from which jets are reconstructed have already been adjusted to point at the event's primary hard-scatter vertex.



(a) (b)

Figure 3.14: Jet energy response (a) and jet mass response (b) as a function of η of the jet.

This alternative mass measurement has better resolution for high-p_T jets with low values of m/p_T. A weighted least-squares combination of the mass measurements is subsequently performed with weights:

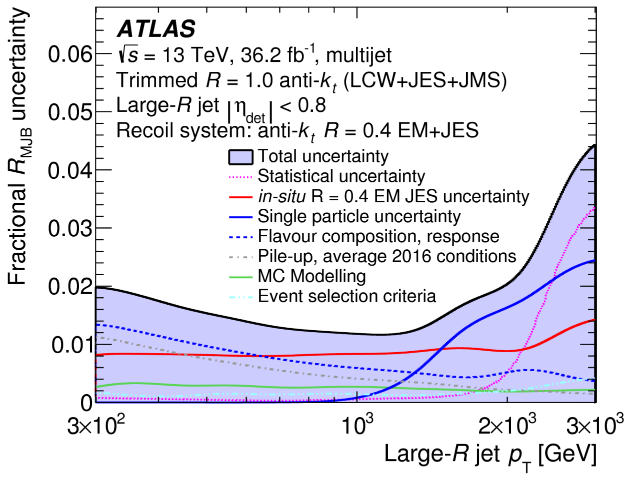
$$m^{comb} = w_{calo} m^{calo} + w_{TA} m^{TA} \tag{3.6}$$

where the two weights, w_{TA} and w_{calo}, are determined considering the expected mass resolution of the two sub-detectors estimation.

3.3.3 Vector Boson Tagging

In addition to the window cut on the jet mass distribution, we use the jet substructure variable, D_2, reconstructed by the energy correlation functions based on energies and pair-wise angles of the sub-constituents[80, 81].

$$D_2^{(\beta=1)} = E_{CF3} \left(\frac{E_{CF1}}{E_{CF2}} \right)^3 \tag{3.7}$$



(a)

Figure 3.15: Fraction uncertainties on the double ratio between data and simulation of the p_T balance obtained in a di-jet process.

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

$$\begin{aligned} E_{CF1} &= \sum_i p_{T,i} \\ E_{CF2} &= \sum_{ij} p_{T,i} p_{T,j} \Delta R_{ij} \\ E_{CF3} &= \sum_{ijk} p_{T,i} p_{T,j} p_{T,k} \Delta R_{ij} \Delta R_{jk} \Delta R_{ki} \end{aligned} \tag{3.8}$$

The tracking detector has excellent angular resolution and good reconstruction efficiency at very high energy, while its energy resolution deteriorates. By combining information from the calorimeter and tracking detectors, the precision of jet substructure techniques can be improved for a wide range of energies. The Track-Calo Clusters (TCCs) are used to reconstruct the large-R jets in this thesis. This procedure is a type of particle flow, complementary to the energy subtraction algorithm as described before. The two algorithms are designed to improve the jet reconstruction performance in very different energy regimes, and to improve the different aspects of jet performance, reflected in their distinct four-vector construction and energy sharing procedures. Energy sharing in the TCC approach is addressed solely based on a weighting scheme where only the relative track momenta are used to spatially redistribute the energy measured in the calorimeter. In practice, this means that the TCC algorithm uses the spatial coordinates of the tracker and the energy scale of the calorimeter. Figure 3.16 shows the comparison of TCC jet mass and D2 resolutions with LCTopo jets used in the previous analysis, as a function of p_T.

A two-dimensional optimization of the jet mass and D_2 thresholds is performed to provide a maximum sensitivity. The cut on jet mass is a window and for D_2 it is one-sided cut. Figure 3.17 shows the optimized thresholds on D_2 and jet mass as a function of p_T.

The 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. The pileup dependency is found to be small. TCC jet tagger achieved higher signal efficiency while keeping similar background rejection to LCTopo jet tagger at highest-p_T region.

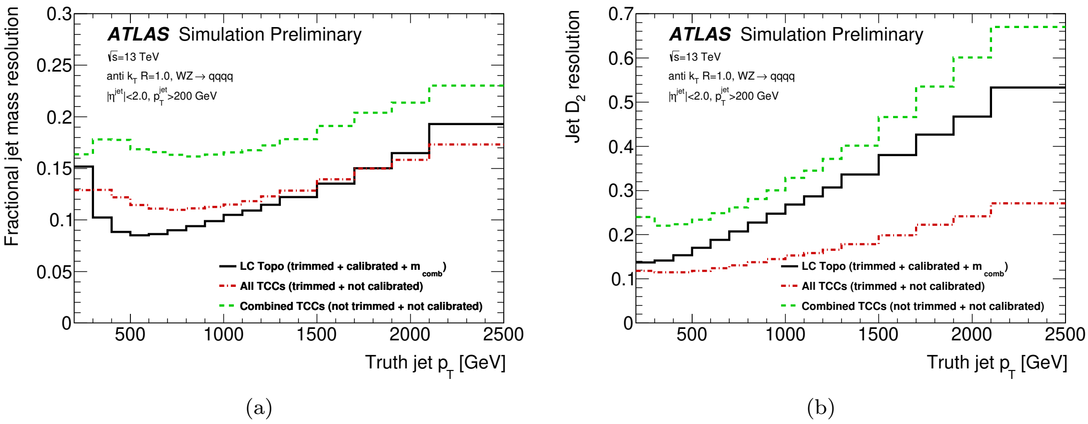


Figure 3.16: A comparison of the fractional jet mass (a) and D_2 (b) resolution for topo-cluster jets (solid black lines), jets built using all-TCC objects (dash-dotted red lines) and jets built using only combined-TCC objects (dashed green lines), as a function of truth-jet p_T [82].

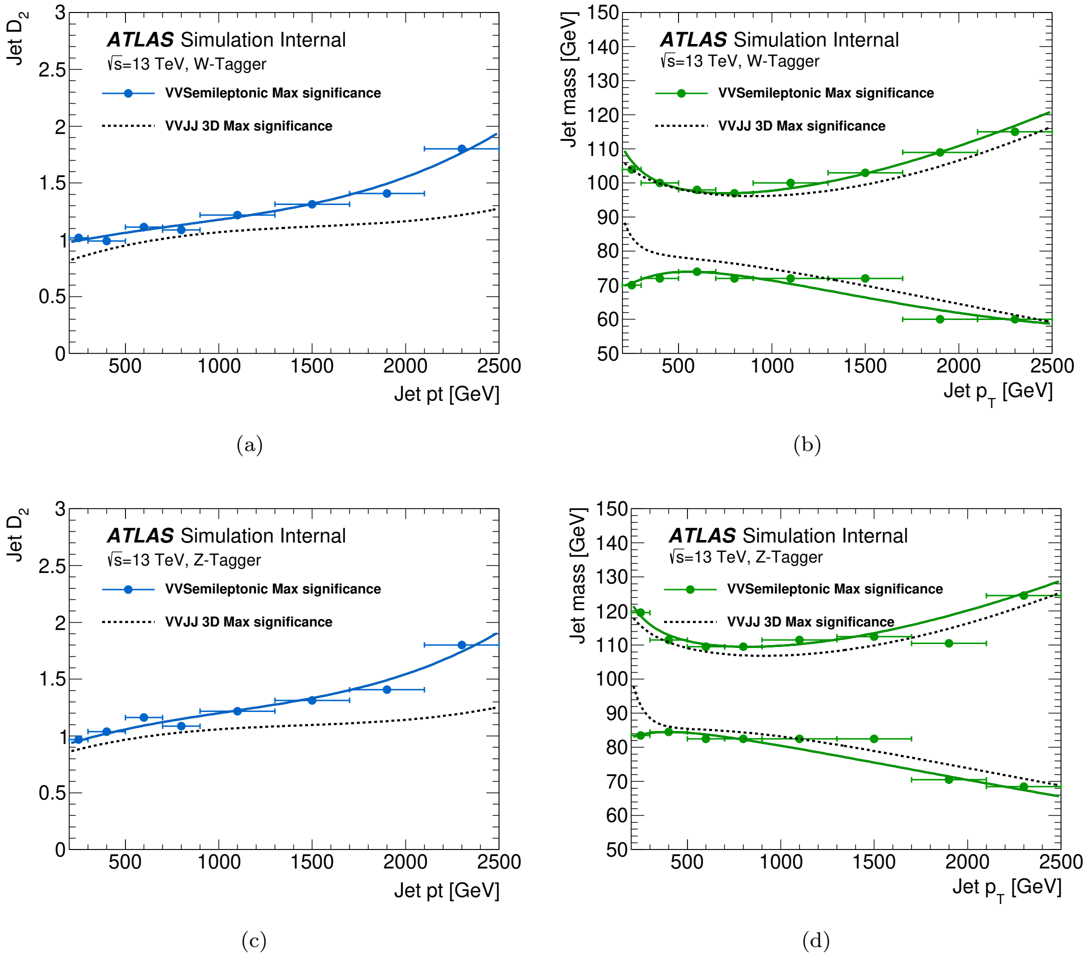


Figure 3.17: The upper cut on D_2 (a) and jet mass window cut i.e. the upper and lower boundary of the mass (b) of the W-tagger as a function of jet p_T. Corresponding values for Z-tagger are shown in (c) and (d). The optimal cut values for maximum significance are shown as solid markers and the fitted function as solid lines. Working points from $VV \to JJ$[83] is also shown as dashed lines as a reference.

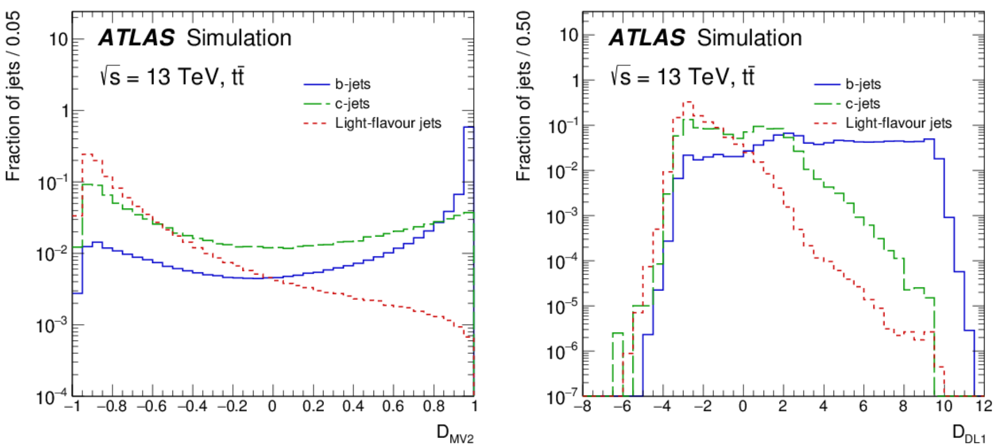
3.3.4 Flavour tagging

The identification of jets containing b-hadrons (b-jets) against the large jet background containing c-hadrons but no b-hadron (c-jets) or containing neither b- or c-hadrons (light-flavour jets) is of major importance in many areas of the physics programme of the ATLAS experiment.

The ATLAS Collaboration uses various algorithms to identify b-jets [84], referred to as b-tagging algorithms, when analysing data recorded during Run 2 of the LHC (2015-2018). These algorithms exploit the long lifetime, high mass and high decay multiplicity of b-hadrons as well as the properties of the b-quark fragmentation. Measurable b-hadrons have a significant mean flight length in the detector before decaying, generally leading to at least one vertex displaced from the hard-scatter collision point. The strategy developed by the ATLAS Collaboration is based on a two-stage approach. Firstly, low-level algorithms reconstruct the characteristic features of the b-jets via two complementary approaches, one that uses the individual properties of charged-particle tracks, associated with a hadronic jet and a second which combines the tracks to explicitly reconstruct displaced vertices. These algorithms, first introduced during Run-1 [84] , have been improved and retuned for the Run-2 [85]. Secondly, the results of the low-level b-tagging algorithms are combined in high-level algorithms based on multivariate classifiers in order to maximise the b-tagging performances. Low-level b-tagging algorithms fall into two broad categories. A first approach, IP2D and IP3D algorithms, is inclusive while the second one explicitly reconstructs displaced vertices (SV1 and JetFitter):

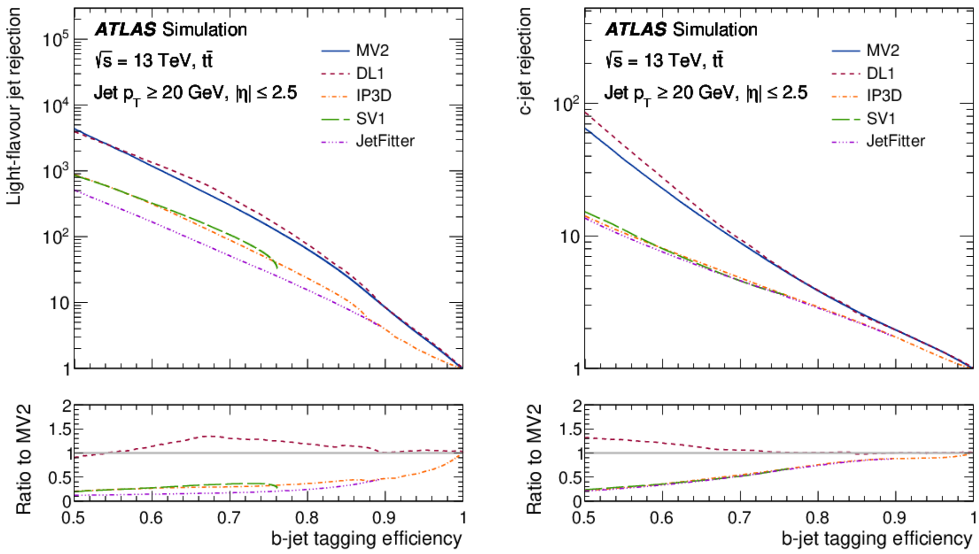
- IP2D and IP3D [86] exploits the large impact parameters of the tracks originating from the b-hadron decay;
- SV1 [87] reconstruct an inclusive secondary vertex;
- JetFitter [88]: reconstruct the full b-to-c-hadron decay chain.

To maximise the b-tagging performance, low-level algorithm results are combined using multivariate classifiers. To this end, two high-level tagging algorithms have been developed. The first one, MV2 [86], is based on a boosted decision tree (BDT) discriminant, while the second one, DL1 [86], is based on a deep feed-forward neural network (NN). The output shapes of the two high-level classifiers are showed in Figure 3.18 and the performances of both low and high-level classifiers in Figure 3.19.



(a) (b)

Figure 3.18: Distribution of the output discriminant of the (a) MV2 and (b) DL1 b-tagging algorithms for b-jets, c-jets and light-flavour jets in the baseline $t\bar{t}$ simulated events.

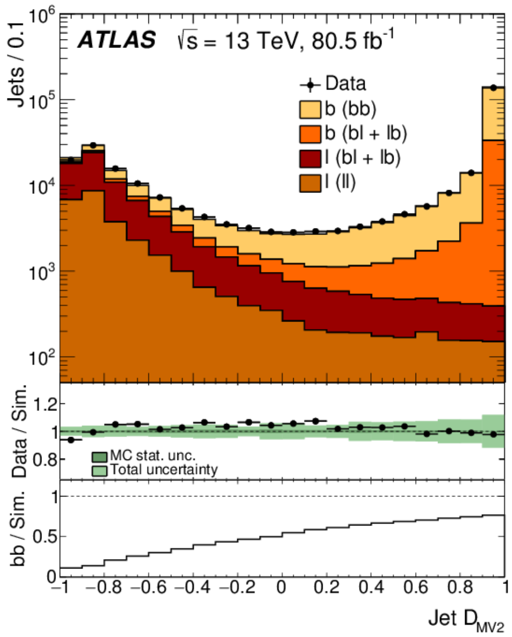


(a) (b)

Figure 3.19: The (a) light-flavour jet and (b) c-jet rejections versus the b-jet tagging efficiency for the IP3D, SV1, JetFitter, MV2 and DL1 b-tagging algorithms evaluated on the baseline $t\bar{t}$ events.

In this thesis work the b-tagging algorithm used is the MV2, even if the performances of the DL1 approach would be interesting for future improvements. The MV2 algorithm consists of a boosted decision tree (BDT) algorithm that combines the outputs of the low-level tagging algorithms described. The BDT algorithm is trained using the ROOT Toolkit for Multivariate Data Analysis (TMVA) [89] on the hybrid $t\bar{t}$ sample. The kinematic properties of the jets, namely p_T and $|\eta|$, are included in the training in order to take advantage of the correlations with the other input variables. However, to avoid differences in the kinematic distributions of signal (b-jets) and background (c-jets and light-flavour jets) being used to discriminate between the different jet flavours, the b-jets and c-jets are reweighted in p_T and $|\eta|$ to match the spectrum of the light-flavour jets. For training, the c-jet fraction in the background sample is set to 7%, with the remainder composed of light-flavour jets. This allows the charm rejection to be enhanced whilst preserving a high light-flavour jet rejection. The BDT training hyper-parameters have been optimised to provide the best separation power between the signal an the background. The output discriminant of the MV2 algorithm for b-jets,c-jets and light-flavour jets evaluated with the baseline $t\bar{t}$ simulated events are shown in Figure 3.20.

The identification of b-jets is used in the resonant search presented in this thesis. The b-tagging algorithm is applied both in the regime in which the boson decay is reconstructed as two small-R jets both in the regime in which one large-R jets is reconstructed; Variable-radius (VR) jets are used to identify large-R jets containing b-hadrons, this is applied for the first time in ATLAS. They are reconstructed using the anti-k_t algorithm from ID tracks associated with large-R jets with a p_T-dependent radius R between 0.02 and 0.4 and a ρ-parameter of 30 GeV [90], and are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. The same b-tagging algorithm for small-Rjets is applied to identify variable-radius jets from b-hadrons.



(a)

Figure 3.20: Distribution of the output discriminant of the MV2 algorithm for the jets in events. Simulated events are classified according to the flavour composition of the two jets, where the first term in each legend entry represents the flavour of the jet which is plotted (b or l) and the following term in parenthesis represents the flavour composition of the event (bb, bl+lb or ll). The ratio panels show the data-to-simulation ratio as well as the fraction of bb events among the simulated events.

3.3.5 Quark/Gluon jet identification

Jet properties can be used to differentiate quark-initiated jets from gluon-initiated jets and in principle could help in signal to background discrimination in analyses that expect only quark-jets in the final state, such as the (VBS/VBF) diboson analyses. A dedicated selection or, in general, a *Quark/Gluon Tagger* could improve the sensitivities of these kind of analyses.

As expected from the theory, the gluon-jets are characterized by a greater number of charged particles with a larger spread respect to the quark-jets because the gluons have a double colour charge respect to the single charge of the quarks. This represents a key element in the q/g discrimination. A simplified Tagger based on the number of tracks inside the jet has been implemented in ATLAS [91]; the charged particles multiplicity represents a discriminating variable, anyway more complex tagger based on more of one single variable are under study [92].

Figure 3.21 shows the distribution of the jet reconstructed track multiplicity (n_{Tracks}) in different p_T ranges with the Pythia 8 generator and processes with a full simulation of the ATLAS detector. Jets must be fully within the tracking acceptance ($|\eta| < 2.1$) and tracks are required to have $p_T > 500\ MeV$ and pass additional quality criteria [91].

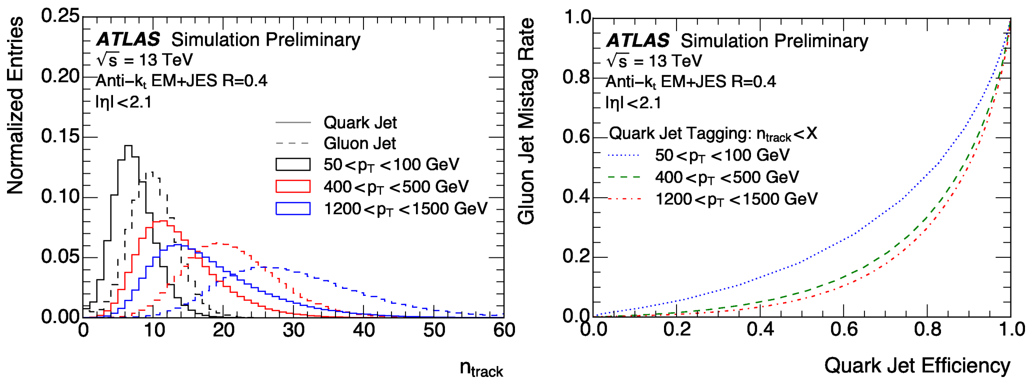


Figure 3.21: Distribution of the jet reconstructed track multiplicity (nTrack) in different p_T ranges (a) and gluon mis-tagging rate as a function of the quark efficiency selection for different p_T ranges.

Figure **??**-(a) shows the gluon jet mis-tagging rate as a function of the quark jet efficiency using the jet track multiplicity as discriminant, with the same simulation setup as in Figure 3.21-(b). Each line corresponds to jets in different p_T ranges. Each curve was generated by placing an upper threshold on n_{Tracks}. A threshold of 0 corresponds to zero efficiency for both quark and gluon jets while a threshold of ∞ corresponds to 100% for both types of jets.

The quark/gluon identification in ATLAS is going in the direction of MultiVariates approach, in particular, a BDT-based identification is under development using the tracks multiplicity, the width and sub-structure information of the jets as discriminating variables.

3.4 Missing energy

In experiments at the colliders, conservation of momentum in the plane transverse to the beam axis implies that the transverse momentum of the collision products should sum to zero. Any imbalance is known as "missing transverse momentum", or E_T^{miss}, and may be indicative of weakly-interacting, stable particles in the final state. Within the Standard Model, this arises from neutrinos.

In the following, the E_T^{miss} performance is studied in two complementary topologies with and without genuine E_T^{miss}, $W \to e\nu$ and $Z \to \mu\mu$, using the first data of the Run-II data taking corresponding to an integrated luminosity of 6 pb^{-1} [93]; further, also the agreement with the MC simulation is tested. Selected calibrated hard objects are used to measure the missing transverse momentum in an event for the E_T^{miss} reconstruction. The $E_{x(y)}^{miss}$ components are evaluated as:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,soft} \tag{3.9}$$

where each object term is given by the negative vectorial sum of the momenta of the respective calibrated objects. Calorimeter information are associated with these reconstructed objects: electrons (e), photons (γ), hadronically decaying τ-leptons, jets and muons (μ). The soft term is reconstructed from detector signal objects not associated with any previous object. These can be ID tracks (track-based soft term) or calorimeter signals (calorimeter-based soft term). Then, the following variables are derived:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \qquad and \qquad \Phi^{miss} = \arctan\left(E_y^{miss}/E_x^{miss}\right) \tag{3.10}$$

The performance of E_T^{miss} reconstruction in $Z \to \mu\mu$ and $W \to e\mu$ data events is compared with the expected performance from the MC samples for the signal and the relevant background processes (Figure 3.22); genuine E_T^{miss} due to neutrino is expected in the second process contrary to the first one in which the fake E_T^{miss} contribution is due to interacting SM particles which escape the acceptance of the detector, are badly reconstructed or fail to be reconstructed altogether.

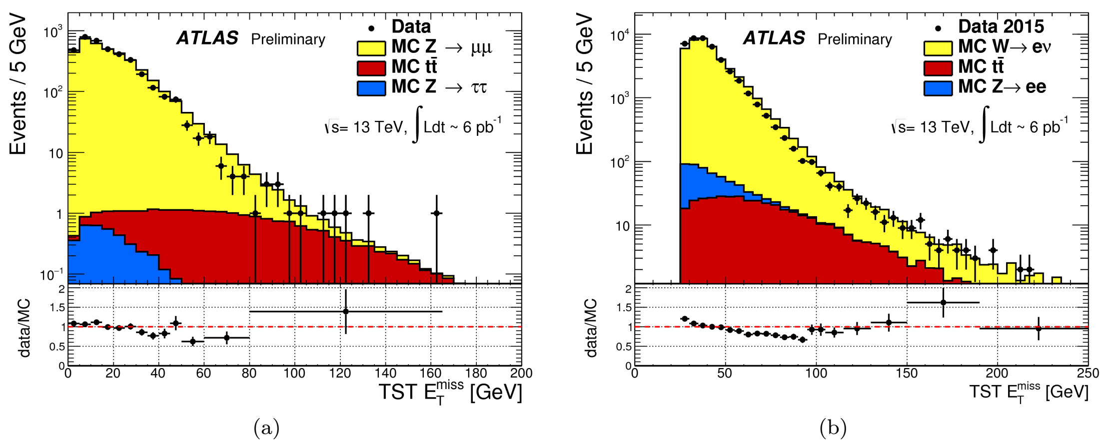


Figure 3.22: Distributions of E_T^{miss} in $Z \to \mu\mu$ events (a) and $W \to e\mu$ events (b). The expectation from MC simulation is superimposed and normalized to data, after each MC sample is weighted with its corresponding cross-section.

The performances of the E_T^{miss} reconstruction is showed evaluating the standard deviation of the E_T^{miss} distribution as a function of the p_T of the hard objects reconstructed in the event (Figure 3.23), in particular, the E_T^{miss} is divided in the two component, longitudinal and transverse, to the p_T^{hard}; different MC simulation performances are tested comparing to the data.

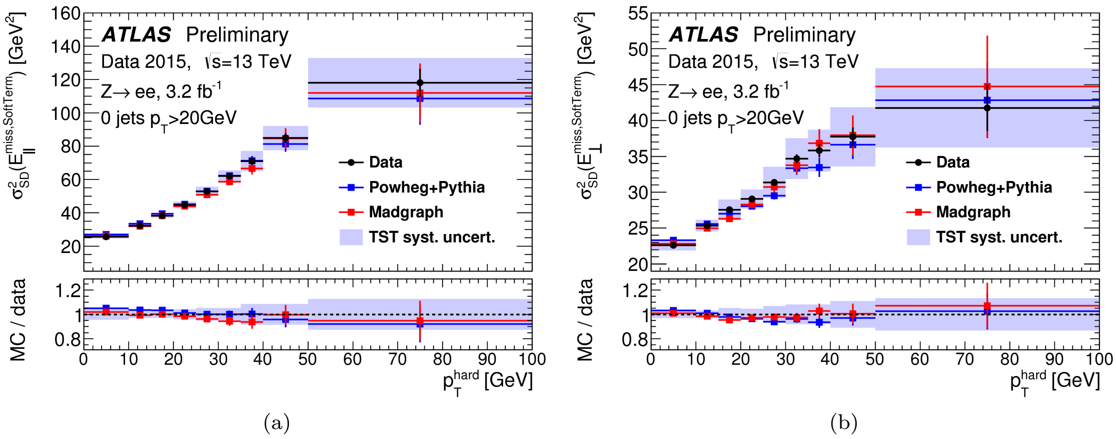


Figure 3.23: Points show the squared standard deviation of the E_T^{miss} soft term distribution projected in the direction longitudinal (a) and in the transverse (b) to p_T^{hard} for $Z \to ee$ with 0-jet events, with vertical bars indicating the statistical uncertainty on the mean. The black points show 25 ns data taken during 2015. Other points show Powheg+Pythia8 and Madgraph Monte-Carlo simulations.

Chapter 4

Machine Learning and Deep learning

In this chapter the main concepts of Machine Learning will be introduced. The Machine Learning (ML) has been one of the main focuses of this thesis; the understanding of the basic principle of a Machine Learning application and of more complex algorithms, as Recurrent Neural Network (RNN), has been crucial in the developing of the physics application that will be illustrated in the chapter 5.6.

After a general overview of the Machine Learning, the attention will move to neural networks and deep learning algorithms with the aim of go trough all the steps needed to develop a model for a classification problem, data pre-processing, learning process up to the features of the RNN.

4.1 What is Machine Learning and its applications

In general, computer or "machines" need to be told what to do; they are strict logic machines with zero common sense. Instead, humans learn from past experiences, or at lest they try; this is the main difference between humans and machines.

This means that if we want machines to do something we have to provide them instructions with step-by-step detailed instructions on what exactly to do. This is the reason fro which we wrote scripts and programs so that machines can follow the instructions. This is the point in which Machine learning arises; the aim is to let computers learn from past experiences.

The machine learning is an application of Artificial Intelligence (AI) that allows computer to get ability to learn automatically and improve their skills without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

The first step of the learning process is based on observations or data, as examples, experiences or instructions, so that the machines could look for patterns in the data and make better decisions in the future based on the experience got with the examples we provide them. The main aim is to allow computer to learn adjusting the actions accordingly during a learning phase, and let the machine without any human intervention after that phase.

4.1.1 Examples of Machine Learning

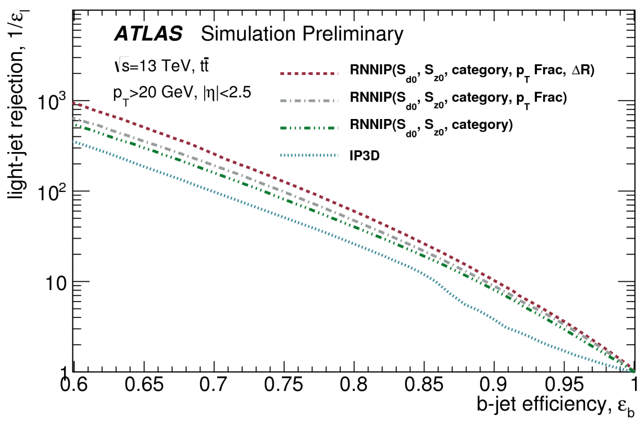
Machine learning is being used in a wide range of applications today. One of the most well-known examples is Facebook's News Feed. The News Feed uses machine learning to personalize each member's feed. If a member frequently stops scrolling to read or like a particular friend's posts, the News Feed will start to show more of that friend's activity earlier in the feed.

Behind the scenes, the software is simply using statistical analysis and predictive analytics to identify patterns in the user's data and use those patterns to populate the News Feed. Should the member no longer stop to read, like or comment on the friend's posts, that new data will be included in the data set and the News Feed will adjust accordingly.

Machine learning is also entering an array of enterprise applications. Customer relationship management (CRM) systems use learning models to analyze email and prompt sales team members to respond to the most important messages first.

More advanced systems can even recommend potentially effective responses. Business intelligence (BI) and analytics vendors use machine learning in their software to help users automatically identify potentially important data points.

Human resource (HR) systems use learning models to identify characteristics of effective employees and rely on this knowledge to find the best applicants for open positions.



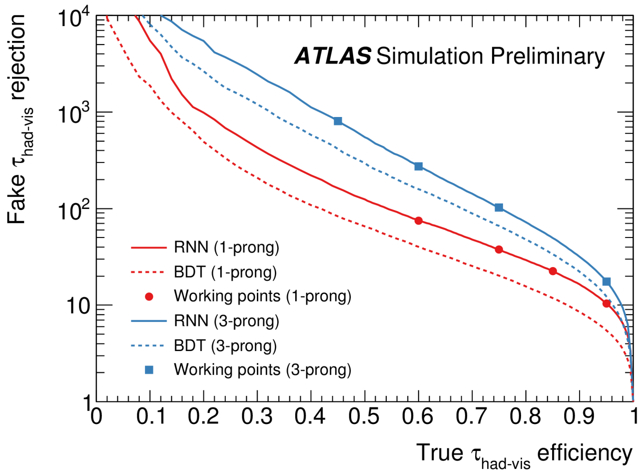
(a)

Figure 4.1: The light-jet rejection versus b-tagging efficiency for jets with $p_T > 20\ GeV$ and $|\eta| < 2.5$, for RNNs trained using various sets of input variables, and for IP3D. The RNN without p_T frac and $\Delta R(track, jet)$ uses only the inputs available to IP3D.

In the high-energy physics the ML are also exploited in order to improve the performances of typical classification problems. In the ATLAS collaboration several application are under development, some for signal to background classification with high-level algorithms as BDT, others that exploit more fundamental problem as the identification of jets. In particular, two RNN-based model are under study; one for the b-jets identification, problem introduced in Section 3.3.4, and the other for the tau-jets identification [94].

Concerning the b-jets identification a novel algorithm is constructed with a Recurrent Neural Network (RNN).This processes charged particle tracks associated to jets without reliance on secondary vertex finding, and can augment existing secondary-vertex based taggers. In contrast to traditional impact-parameter-based b-tagging algorithms which assume that tracks associated to jets are independent from each other, the RNN based b-tagging algorithm can exploit the spatial and kinematic correlations between tracks which are initiated from the same b-hadrons [95]. Better performances respect to the IP3D algorithm are reached.

The algorithm RNN-based for the τ-jets identification [96] employs information from reconstructed charged-particle tracks and clusters of energy in the calorimeter associated to $\tau_{had-vis}$ candidates as well as high-level discriminating variables. Better performances respect to BDT using only high-level variables are showed in Figure 4.2.



(a)

Figure 4.2: Rejection power for quark and gluon jets misidentified as $\tau_{had-vis}$ (fake $\tau_{had-vis}$) depending on the true $\tau_{had-vis}$ efficiency. The curves for 1-prong (red) and 3-prong (blue) $\tau_{had-vis}$ candidates using the RNN-based (full line) and the BDT-based (dashed line) identification algorithms are shown.

4.2 Machine Learning techniques

Usually, the machine learning algorithms are categorised according to the characteristic of the training phase.

- Supervised Machine Learning algorithms: Labeled examples are used during the learning phase, future events will be classified according the kind of labeled used. The analysis of the known and labeled dataset let the learning algorithm to produce an inferred function to make predictions about the output values.
- Unsupervised Machine Learning algorithms: The informations used during the training phase are neither classified or labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data.
- Semi-supervised Machine Learning algorithms: Semi-supervised machine learning algorithms fall somewhere in between supervised and unsupervised learning since they use both labeled and unlabeled data for training, typically a small amount of labeled data and a large amount of unlabeled data. This method allows in general to improve the accuracy of the future prediction respect to the two previous categories.
- Reinforcement machine learning algorithms: Reinforcement machine learning algorithm is a learning method that interacts with its environment by producing actions and discovers errors or rewards. This method allows machines and software agents to automatically determine the ideal behaviour within a specific context in order to maximize its performance.

4.2.1 Deep Learning

Deep Learning is a sub-field of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks [97].

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the back-propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech [98]. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned.

4.2.2 Recurrent Neural Networks

The general structure of the neuron used for deep feed-forward networks are a special case. Indeed, all the connections goes in one way and only to the layer immediately on the right (Figure 4.7); so, the update rule of the $i-th$ neuron in the l layer is:

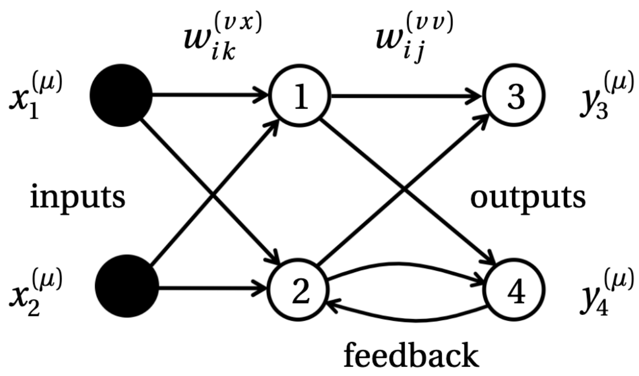
$$V_i^l = f\Big(\sum_j w_{ij}^l V_j^{l-1}\Big) \tag{4.1}$$

The back-propagation algorithm relies on this feed-forward layout. It means that the derivatives

$$\frac{\partial V_j^{l-1}}{\partial w_{mn}^l} = 0 \tag{4.2}$$

vanish. This ensured the simple iterative structure of the back-propagation.

In same cases could be useful or needed to not have this simple layout. More general networks have a feed-forward layout with feedbacks, as in Figure 4.3.



(a)

Figure 4.3: Network with a feedback connection.

These kind of networks are called Recurrent Neural Networks. There are many different ways in which the feedbacks can work: from the output layer to hidden neurons for example or there could be connections between the neurons in a given layer.

Recurrent networks can be used to learn sequential inputs, as in speech recognition and machine translation. The informations needed for the training set are a time sequence of input features and a sequence of targets $[x(t), \vec{y(\vec{y})}]$, in general, the two sequences have different dimensions. The job is to train the network over the time sequence of inputs to predict the targets.

The variable t is related to the time step, so it is discrete, and the network will be made of recurrent layer(s) and of a output layer, us usual; so, the updates rules are:

$$V_i(t) = f\Big(\sum_j w_{ij}^{VV} V_j(t-1) + \sum_k w_{ik}^{Vx} x_k(t) - \theta_i^V\Big) \tag{4.3}$$

$$O_i(t) = g\Big(\sum_j w_{ij}^{OV} V_j(t) - \theta_i^O\Big) \tag{4.4}$$

in general, the activation functions of the recurrent layer, f, and the one of the output layer, g, could be different.

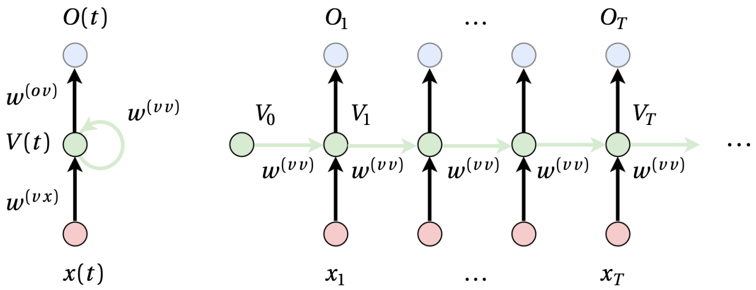
The back-propagation through time is used to train recurrent networks with time-dependent inputs and targets. The idea is to unfold the network in time to get rid of the feedbacks, at the expense of as many copies of the original neurons as there are time steps (Figure 4.4). The unfolded network has T inputs and outputs. It can be trained in the usual way with stochastic gradient descent. The errors are calculated using back-propagation, but here the error is propagated back in time, not from layer to layer.

Recurrent networks are used for machine translation [99].

The general vanishing gradient problem https://towardsdatascience.com/the-vanishing-gradient-problem-69bf08b15484 is solved improving the architecture of the network.

The idea is to replace the hidden neurons with computation units (cells) that are especially designed to solve the vanishing gradient descent problem; one kind of this cells are the basis, for instance, of the the Long Short Term Memory (LSTM) method [100], that is the architecture used in this thesis 5.6.

For machine translation, it is possible to represent words in a dictionary so that, for example, 100... represents the first word in the dictionary, 010... the second word, and so forth. Each input is a vector with as many components as there are words in the dictionary. A sentence corresponds to a sequence $x_1, x_2, ..., x_T$. Each sentence ends with an end-of-sentence tag, $< EOS >$. An activation function outputs give the probability $p(O_1, ..., O_{T'}|x_1, ..., x_T)$ of an output sequence conditional on



(a)

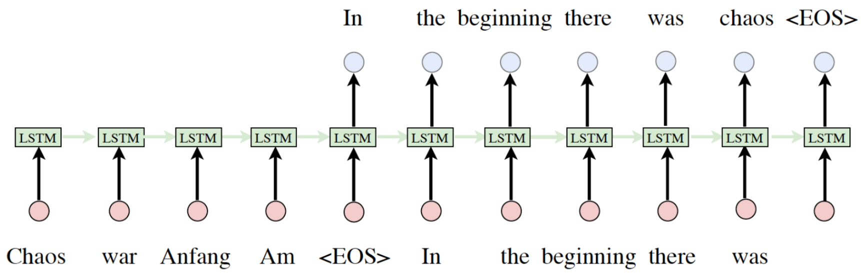
Figure 4.4: Recurrent network with one hidden neuron (green) and one output neuron (blue). The input terminal is drawn red (left). The same network but unfolded in time (right).

the input sequence. The translated sentence is the one with the highest probability (it also contains the end-of-sentence tag<EOS>) [101].

In this way, the network encodes the input sequence $x_1, x_2, ..., x_T$ in these states. Upon encountering the<EOS>tag in the input sequence, the network outputs the first word of the translated sentence using the information about the input sequence stored in the hidden (LSTM) states as shown in Figure 4.5. The first output is fed into the next input, and the network continues to translate until it produces an $< EOS >$ tag for the output sequence. In other terms, the network calculates the probabilities:

$$p(O_1, ..., O_{T'}|x_1, ..., x_T) = \prod_{t=1}^{T'} p(O_T|O_1, ..., O_{T-1}; x_1, ..., x_T) \tag{4.5}$$

where $p(O_T|O_1, ..., O_{T-1}; x_1, ..., x_T)$ is the probability to get the next word in the output sequence given the inputs and the output sequence up to O_{t-1} [102].



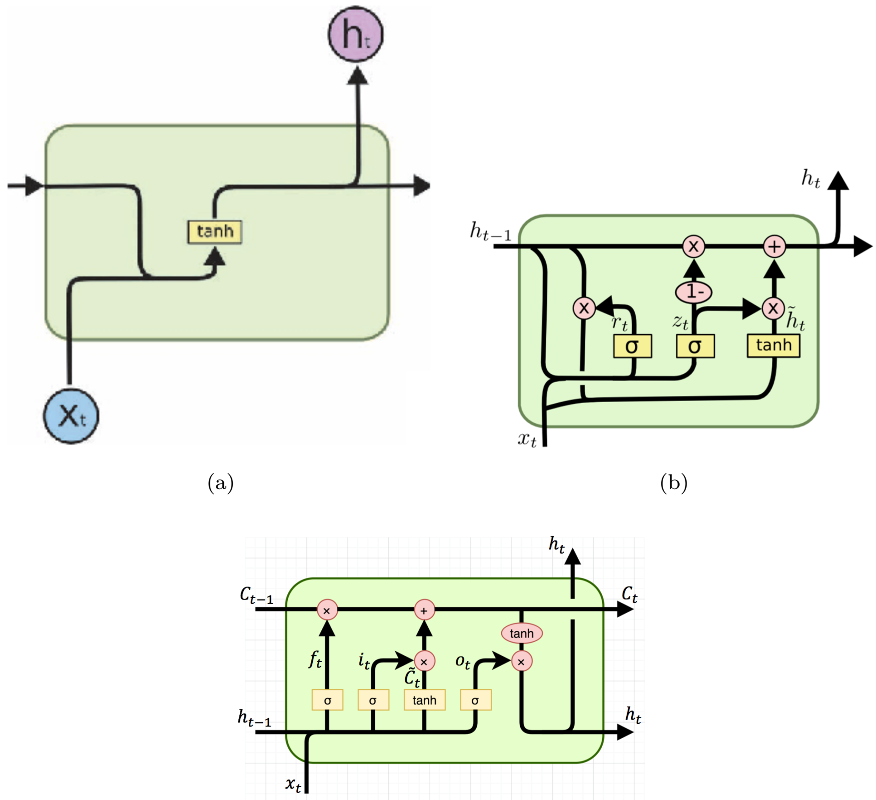
(a)

Figure 4.5: Schematic illustration of unfolded recurrent network for machine translation. The green rectangular boxes represent the hidden states in theform of long short term memory units (LSTM).

Several RNN cell have been developed. The three most common are schematically described:

- simple RNN: there is simple multiplication of Input (x_t) and Previous Output (h_{t-1}). Passed through $tanh$ activation function. No Gates present.
- Gated Recurrent Unit (GRU): an update gate is introduced to decide whether to pass Previous O/P (h_{t-1}) to next Cell (as h_t) or not. Forget gate is nothing but additional Mathematical Operations with a new set of Weights (W_t).
- Long Short Term Memory Unit (LSTM): 2 more Gates are introduced (Forget and Output) in addition to Update gate of GRU. These are additional mathematical operations on same

inputs (x_t and h_{t-1}). So overall, LSTM has introduced 2 mathematical operations having 2 new sets of Weights.



(a) (b)

(c)

Figure 4.6: Scheme of the inner architecture of a RNN cell, SimpleRNN (a), GRU (b) and LSTM (c).

4.3 Implementation of a Machine Learning classifier

Nowdays, several frameworks provides the algorithms of machine learning techniques, and in particular of deep learning networks, encouraging the use of these approaches in more and more sector, as big data analyses, finance, physics. In this thesis, the deep learning networks used have implemented using the Keras libraries [103] and python language [104]. The Keras libraries provide the implementation of the most used layers [105], the user could build the specific architecture that is needed for the problem and, then, the network is trained with a learning algorithm [106].

4.3.1 Data pre-processing

The input dataset for a deep learning approach is usually managed before to pass it into the network.

It is common to first divide the input dataset into a training/test sub-datasets in order to have two independent dataset representative of the same population under study; further, a shuffling of the entire dataset before the splitting could be useful in order to guarantee that the two splitted dataset are good representatives.

Usually, the training dataset is further splitted in a training and a validation dataset, used to monitor the learning phase, as it will explained in the next Section 4.3.4.

The input features are usually scaled before to be processed by the network; usually, two method could be used:

- standard scaling: $x' = \frac{x-\mu}{\sigma}$
- minimum-maximum scaling: $x' = \frac{x-x_{min}}{x_{max}-x_{min}}$

where μ and σ are the mean and the standard deviation of distribution of the input feature x and x_{min} and x_{max} are the minimum and the maximum values of the distribution.

The main point of the features scaling is that all the features that could have numerical range completely different are transformed to the same scale so that the network could give the same initial importance to all of them. The features scaling shows learning processes converging faster and also better performances.

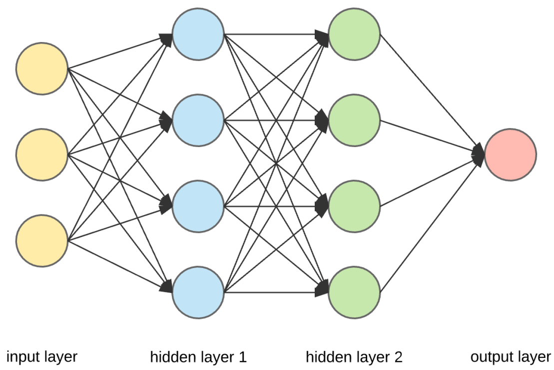
Before to start the learning process the targets of each instance of the dataset has to be fixed. In a supervised problem, the target represents the truth information we would like the model learns. In a binary classification problem, for instance, the model will have just one target and it will be $0, 1$ according to the class, 0 or 1, the single instance belongs to. At this point the input dataset is ready for the learning process.

4.3.2 Learning process

Let consider a neural network; each layer is made of perceptrons or neurons connected to each other; each connection is associated with a weight that controls the importance of that particular relationship in the neuron when the input value arrives.

Each neuron has an activation function that defines the output of the neuron. The activation function is used to introduce non-linearity in the modeling capabilities of the network.

Let consider a set $X = x_1, x_2, ..., x_N$ of input features and a set $Y = y_1, ..., y_M$ targets (this is an example of supervised learning for classification problem).



(a)

Figure 4.7: Scheme of a neural network; the input features are the input of the first layer, the number of hidden layers used impacts on the deep of the network; finally, the output outcome from the last layer.

Training our neural network, that is, learning the values of our parameters (weights w_{ij} and b_j biases) is the most genuine part of Deep Learning and we can see this learning process in a neural network as an iterative process of "going and return" by the layers of neurons. The "going" is a forwardpropagation of the information and the "return" is a backpropagation of the information.

The output of the $i-th$ neuron is given by:

$$n_i = f_{activation}\left(\sum_{j=1}^{N} w_{ij} n_j\right) \tag{4.6}$$

The first phase of forward-propagation is about the evaluation of the outputs of the network given the input data and passing through each layer and each connection of the network. Once the data crossed all the connection and all the neurons made their transformations the last layer return values that we expect to be the targets we gave in order to label the data.

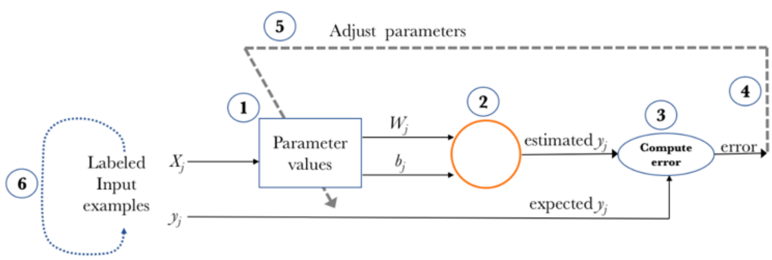
A this point we use a loss function to estimate the loss (or error) and to compare and measure how good/bad our prediction result was in relation to the correct result. Ideally, we want our cost to be zero, that is, without divergence between estimated and expected value. Therefore, as the model is being trained, the weights of the interconnections of the neurons will gradually be adjusted until good predictions are obtained.

Once the loss has been calculated, this information is propagated backwards. Starting from the output layer, that loss information propagates to all the neurons in the hidden layer that contribute directly to the output. However, the neurons of the hidden layer only receive a fraction of the total signal of the loss, based on the relative contribution that each neuron has contributed to the original output. This process is repeated, layer by layer, until all the neurons in the network have received a loss signal that describes their relative contribution to the total loss.

At this point, the weights of the network have to be updated. The goal is to make the loss as close as possible to zero after each iteration and update of the weights. Usually, a gradient descent technique is used. According to this the weight are changed with small variations with the help of the calculation of the derivative (or gradient) of the loss function, which allow to see in which direction in the weights space we should go to reach the global minimum.

The learning process could be summarised as follow (Figure 4.8):

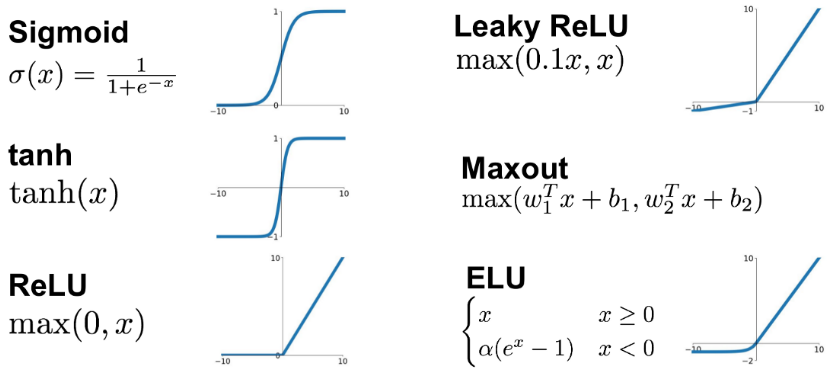
1. Start with values (often random) for the network parameters.
2. Take a set of examples of input data and pass them through the network to obtain their prediction.
3. Compare these predictions obtained with the values of expected labels and calculate the loss with them.
4. Perform the back-propagation in order to propagate this loss to each and every one of the parameters that make up the model of the neural network.
5. Use this propagated information to update the parameters of the neural network with the gradient descent in a way that the total loss is reduced and a better model is obtained.
6. Continue iterating in the previous steps until we consider that we have a good model.



(a)

Figure 4.8: Scheme of the steps of the learning process of a neuron.

The activation function are used to propagate the information of the input data through a neuron; they are used to introduce non linearity in the model. In Figure 4.9 the most used activation functions are showed.



(a)

Figure 4.9: Functional trend of the most used activation functions.

4.3.3 Back-propagation component

Back-propagation could be considered as a method to update the parameters of the network. Once the loss function is evaluated, the parameters will be updated in the reverse orders taking into account this loss information.

The loss function represents the way to evaluated how close the neural network is close to the idea behaviour we would like to reach.

The most used loss function in classification problem is the categorigal crossentropy or the binary crossentropy in the binary classification problems; the function form is given by:

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} y_i \left[log(\hat{y}_i) + (1 - y_i) \cdot log(1 - \hat{y}_i) \right] \tag{4.7}$$

The learning process could be considered as a general optimisation problem in which the parameters of the model must be varied in a way in which the loss function considered is minimised. In general, the solutions to these problem can not be found analytically but they could be searched for with an iterative procedure.

One of the most common optimiser used in deep learning problems is the Gradient Descent, that is also a basis for other methods. In particular, the first derivative (the gradient) of the loss function is used when the parameter are needed to be updated. The gradient are evaluated in the dimension of the specific parameter we are considering in the process and a variation is applied to the parameter, according to:

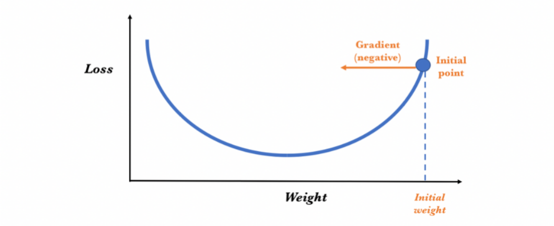
$$\delta w_{ij} = -\alpha \frac{\partial L}{\partial w_{ij}} \tag{4.8}$$

where α is used named learning rate. The idea of gradient descent optimiser is showed in Figures 4.10.

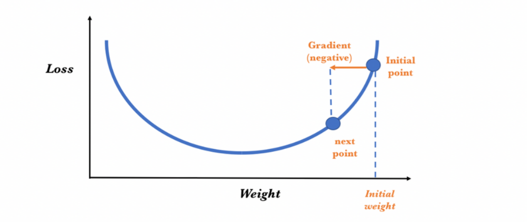
Once the idea of the learning process on one single event is explained, we should consider how use the dataset available for the training has to be used for the full training of the model. In general, the full dataset is used several times, each iteration over the full dataset is called *training epoch*.

According to how many times the parameters are updated, 3 types of gradient descent method arise:

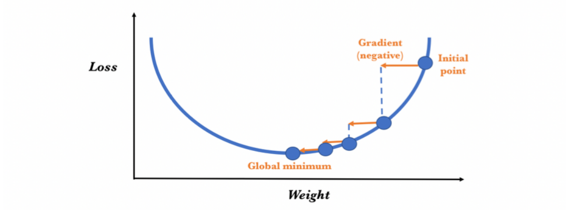
- *Stochastic Gradient Descent* The loss function is evaluated and the parameters are updated after each single event of the training dataset.
- *Batch Gradient Descent* The loss function is evaluated for each event of the training dataset but the model is updated after all the events have been processed, that means after an epoch.



(a)



(b)



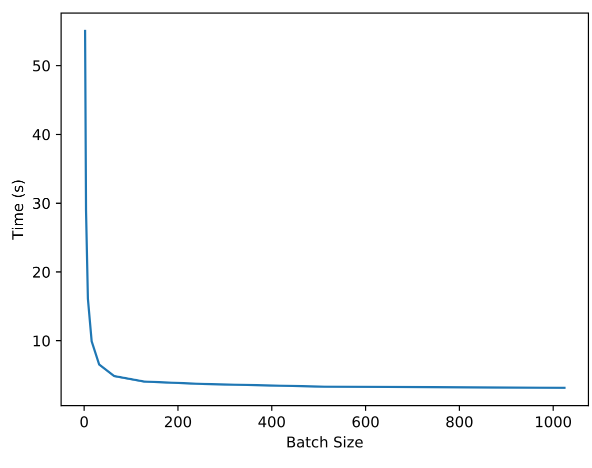
(c)

Figure 4.10: Schematic representation of the gradient descent method.

- *Mini-Batch Gradient Descent* The training dataset is splitted into small batches that are used to evaluate the loss function and to update the model parameters. Implementations may choose to sum the gradient over the mini-batch which further reduces the variance of the gradient.

Mini-batch gradient descent seeks to find a balance between the robustness of stochastic gradient descent and the efficiency of batch gradient descent. It is the most common implementation of gradient descent used in the field of deep learning.

In addition, the simultaneous calculation of the gradient for many input events can benefit of the parallel capability of the machine that is used for the computation (Figure 4.11); this can be done using matrix operations that are implemented very efficiently with modern parallel processors as GPUs.



(a)

Figure 4.11: Compution time needed during a training epoch varying the size of the batch; a machine with 24 parallel core has been used.

4.3.4 Learning Curve and diagnosis

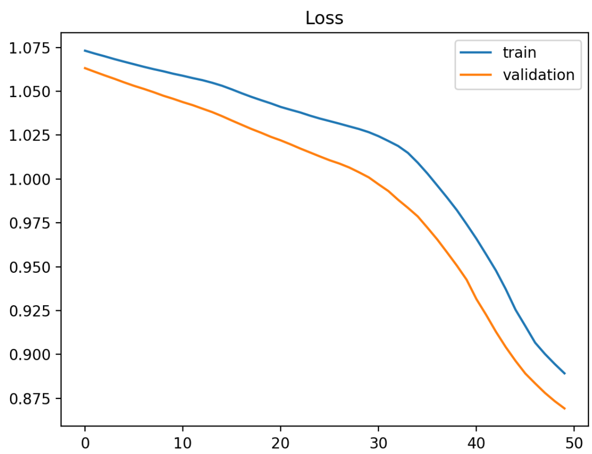
A learning curve is a plot of model learning performance over experience or time.

Learning curves are a widely used diagnostic tool in machine learning for algorithms that learn from a training dataset incrementally. The model can be evaluated on the training dataset and on a hold out validation dataset after each update during training and plots of the measured performance can created to show learning curves.

Reviewing learning curves of models during training can be used to diagnose problems with learning, such as an underfit or overfit model, as well as whether the training and validation datasets are suitably representative.

There are three common dynamics that you are likely to observe in learning curves; they are:

- Under-fitting: the loss function showed in Figure 4.12 decreases slowly during the training phase and it did not reached a plateau yet; more training epochs could give a more accurate model.

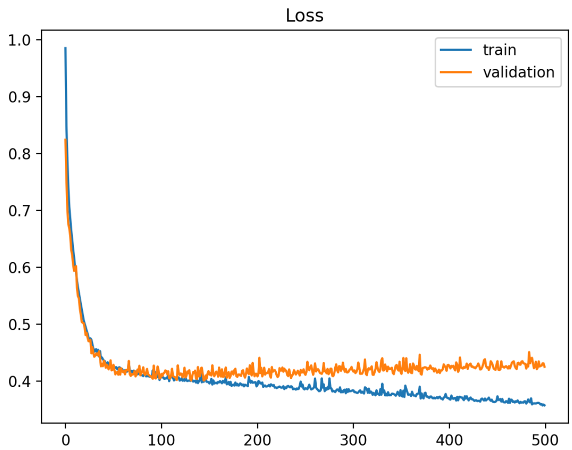


(a)

Figure 4.12: Loss function that shows an under fitting during the learning phase.

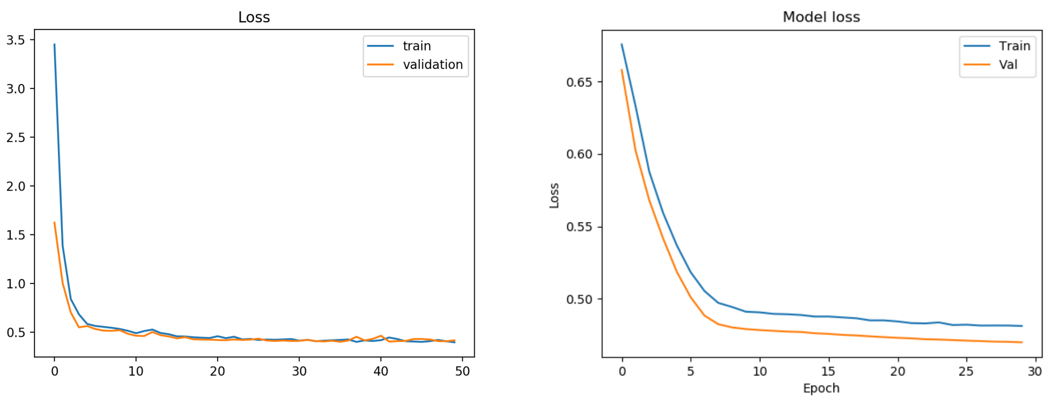
- Over-fitting: the loss function in Figure 4.13 reaches a reasonable (small) value and a plateau but the validation dataset shows that the performances are worst respect to the training dataset starting from an epoch ($\sim$ 100 in this example), in other words the model works better on the training dataset and works worst on a "new" dataset.

- Good fitting: the loss curves showed in Figures 4.14 represent a good model. The loss curve reach a plateau and they do not show any trend to become larger after several training epochs; the validation curves is always comparable and smaller than the training one and it does not never become significantly larger; the distance between the training and validation curve in a good model depends on several factors, the main are the number of training steps in a single epochs (effect of the so called running-mean), the dropout fraction and the relative size between the two dataset used.



(a)

Figure 4.13: Loss function that shows an over fitting during the learning phase.



(a) (b)

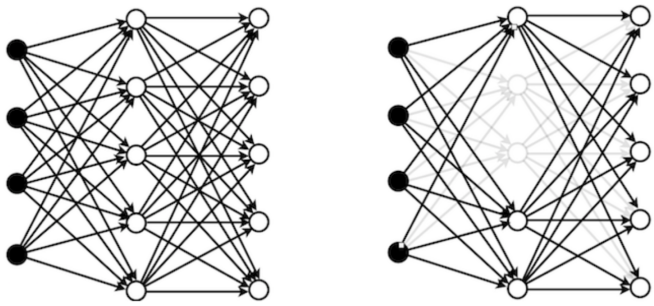
Figure 4.14: Loss function that shows a good trained model.

4.3.5 Regularisation

The problem of the over-fitting of a network become more severe as the network become deeper and deeper; indeed, the number of hidden layers, of neurons and of connections could drastically increase with deep neural networks. Therefore, regularisation schemes, that in general are designed to prevent the tendency to over-fit, are really important for deep networks.

Common regularisation scheme are the L_1 and L_2 [97].

Other new regularisation techniques have been developed to prevent over-training, as drop-out [107]. In this scheme neurons are ignored during the training phase; usually, this is applied only to neurons in the hidden layers (Figure 4.15).



(a)

Figure 4.15: Illustration of the drop-out idea.

The implementation of this scheme foresees that the training algorithm ignores a fraction p of neurons for each hidden layer of the model and, for each step of the training algorithm (for each mini-batch or for each single pattern), update the weight in the remaining network as usual. In this way, the weights of the dropped neurons are not updated as their outputs. For the next step of the training algorithm the neurons dropped are put back and another set corresponding to a fraction p of the total neurons are dropped. Once the training passes trough each step and finished all the neurons are activated again.

The authors of this scheme motivated the idea by the fact that the performances of machine learning algorithms are usually improved when the outputs of several learning attempts are combined. In principle, several learning training varying the setup of the layers could be done and at the end the output could be summed as an average, but this could be more expensive by a computation point of view. The point is that the drop-out corresponds in practise to the case of the training of a large number of different networks. If, for instance, there k neurons in each layer, there are 2^k different combination of neurons that are turned off. The general result is that the network may benefit of the drop-out and learns more robust characteristics of the inputs, reducing the over-fitting.

Chapter 5

Search for new physics in diboson final states

In this chapter, the searches for new physics in diboson final states will be discussed. Firstly, the experimental features and the selection procedure applied in the search for heavy resonances will be given. The description will focus mainly on the new Machine Learning application that has been introduced to improve the categorisation of the production mechanism of the resonance.

Then, the main results of a non-resonant analysis namely the search for the diboson in association with two high energy jets will be presented, as published in the paper [108].

5.1 Analysis Methodology

A physics analysis of the data collected from the high energy collisions at the LHC is usually designed to enhance the rate of the events related to an interesting process under investigation (signal process), if it is predicted by the SM or if it is a new physics process, against the background represented by the other SM processes (background processes). In order to do that, the key kinematics features of the interesting processes are studies and used to reduce the background that could be really different from the signal process (reducible background); this could be done using cut based selection on these discriminating features or with ML approaches. At the end, the phase space selected as to be the most sensitive to the signal process is called Signal Region (SR). Unfortunately, it will not be never possible to reduce the entire background processes with a SR definition, but part of the background will still have not null probability to be in the SR phase space (irreducible background). The goal of the analysis is to have enough sensitivity to be able to directly observe the signal process in the real data that will be in the SR. Several strategies could be used in the constraints of the irreducible background. A Control Region (CR) is defined as a phase region in which no signal processes are expected but to be more similar to the SR phase space as possible; usually, just one cut is reverse respect to the SR definition. The CRs can be used to constraint the behaviour of the SM background expected. In the analyses presented in this chapter, the strategy has been to believe in the description of the background processes coming from MC simulations. In order to better improve the MC modelling the CRs are used to take more informations directly from the data; the MC simulations are improved constraining the simulations to the observed data (fit to CR/SR with $only - Background$ hypothesis). Then, the corrected simulations are used as a baseline where the signal process will manifest itself as an excess over it. The signal process is observed if the statistical significance is evaluated to be $> 5\sigma$. In this case the observed events can be used to measure more informations, such as the cross section of the process or other dynamical features. If the process is not observed, the fit is redone with this time the $Signal + Background$ hypothesis and exclusions limits of the cross section of the signal process are derived.

5.2 Data and MonteCarlo samples

The results reported in this chapter use the pp collision data collected during the period of the full Run2, from 2015 to 2018, corresponding to an integrated luminosity of $139.0 \pm 2.4\ fb^{-1}$ (table 5.1). The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [109], obtained using the LUCID-2 detector [110] for the primary luminosity measurements.

Year	L $[fb^{-1}]$
2015	3.21
2016	32.88
2017	44.31
2018	58.45
Combined	139.0 ± 2.4

Table 5.1: Integral luminosity from 2015 to 2018 used in the analysis.

Monte Carlo simulations are used to generate events needed for background modelling, evaluation of the signal acceptance, optimisation of event selection, estimation of systematic uncertainties and the statistical analysis. They are produced using ATLAS approved event generator settings. The EvtGen v1.2.0 program [111] is used for properties of the bottom and charm hadron decays. Additional pp collisions generated with Pythia 8.186 [112] are overlaid to model the effect of the pileup for all MC events. All samples are processed through the full ATLAS detector simulation [113] based on GEANT4 [114]. All simulated events are processed with the same trigger and reconstruction algorithm as the data.

As discussed in Section 1.4, the main background processes for this analysis are $W/Z + jets$, $t\bar{t}$, singletop, SM diboson and multi-jets productions.

Events containing W or Z bosons with associated jets are simulated using the Sherpa 2.2.1 generator [115] using the NLO matrix elements for porcesses with up to 2 jets and LO for up to 4 jets. The NNPDF3.ONLO set of PDFs has been used and the events are normalised to the NNLO cross sections prediction.

The $t\bar{t}$ and single-top events are generated with the Powheg-Box [116] generator with the NNPDF3.0NLO 513 PDF [117] sets in the matrix element calculation. The top quark spin correlations are preserved (for t-channel, top quarks are decayed using MadSpin [118]). For all processes the parton shower, fragmentation, and the underlying event are simulated using Pythia8.230 with the A14 tune set [119]. The parameter h_{damp} to regulate the high-pT radiation in the Powheg is set to $1.5mt$ for a good data/MC agreement at high pT region [120].

The diboson processes (WW, WZ and ZZ) are generated with Sherpa2.2.1 [115] generator.

The electroweak production of diboson processes (WWjj, WZjj and ZZjj) are neglected due to the extremely small cross sections.

The signal samples for the three types of resonances, described in Section 1.5.1, are generated with MadGraph5 [121] and interfaced with Pythis 8; for the ggF RS Graviton and HVT (both DY and VBF) MadGraph5-2.2.2 interfaced to Pythia 8.186 is used while for the Radion signals (both ggF and VBF) and VBF RS Graviton MadGraph5-2.6.0 interfaced to Pythia 8.212 is used.

The RS Graviton samples are generated with $k/M_{Planck} = 1$ and the radion samples are generated with $k\pi rc = 35$ and $\Lambda_R = 3\ TeV$. The HVT model A signal samples are generated with $g_H = -0.56$ and $g_F = -0.551$. Model B interpretation will be performed assuming the same signal shape as the model A, because the difference on the width in the model B from A is smaller than the detector resolution. 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 6 TeV for each scenario.

5.3 Physics objects selection

The physics objects used in the analysis have been detailed described in Section 3. In this analysis, two different types of leptons have been defined. The EMTopo are used for the small-R jets reconstruction, TrackCaloClusters are used for the large-R jets reconstruction.

Electrons

Two types of electron definition are used in this analysis (Table 5.2).

- "Tight" electron: used to select $W \to e\nu$ candidate in 1-lepton channel.
- "Loose" electron: used to select $Z \to ee$ candidate in 2-lepton channel and to veto events with additional leptons in 0- and 1-lepton channels.

In 1-lepton channel, the electron isolation working points are optimized to minimize the contribution of the non-prompt electrons. In 2-lepton channel, looser isolation working point based on the ID track variable only are preferable to keep high signal efficiency at very high-p_T region, in where two electron candidates are closed by each other. In order to keep a high signal efficiency at high-pT region, we use FCLoose isolation working point only at $p_T < 100\ GeV$ (Section 3.1).

Feature	*Loose*	*Tight*
Pseudorapidity range	$\lvert\eta\rvert < 2.47$	
Transverse momentum	$p_T > 7\ GeV$	$> 30\ GeV$
Track to vertex association	$\lvert d_0^{BL}(\sigma)\rvert < 5$	
	$\lvert\Delta z_0^{BL} \sin\theta\rvert < 0.5$ mm	
Identification	Loose	Tight
Isolation	FCLoose at $p_T < 100\ GeV$ and no isolation requirement at $> 100\ GeV$	FixedCutHighPtCaloOnly

Table 5.2: Summary of electron selection.

Muons

Two types of muon definition are used in this analysis (Table 5.3).

- "Tight" muon: used to select $W \to \mu\nu$ candidate in $l\nu qq$ channel.
- "Loose" muon: used to select $Z \to \mu\mu$ candidate in $llqq$ channel and to veto events with additional leptons in $l\nu qq$ and $\nu\nu qq$ channels.

In 1-lepton channel, the muon identification and isolation working points are optimized to minimize the contribution of the non-prompt muon. In 2-lepton channel, looser working points are chosen to keep high signal efficiency at very high-pT region, in where two muon candidates are closed by each other. The definitions of the Tight and Loose muons are summarized in Table 5.

Overlap Removal

To prevent double-counting coming from leptons and jets a standard ATLAS overlap removal procedure has been applied (Table 5.4)

Criteria	*Loose*	*Tight*
Pseudorapidity range	$\|\eta\| < 2.5$	
Transverse momentum	$p_T > 7$ GeV	$p_T > 30$ GeV
d_0 Significance Cut	$\|d_0^{BL}(\sigma)\| < 3$	
z_0 Cut	$\|z_0^{BL} \sin\theta\| < 0.5$ mm	
Selection Working Point	Loose	Medium
Isolation Working Point	FCLoose at $p_T < 100$ GeV and no isolation requirement at > 100 GeV	FixedCutTightTrackOnly

Table 5.3: Summary of muon selections.

Reject	Against	Criteria
electron	electron	shared track, $p_{T,1} < p_{T,2}$
muon	electron	is calo-muon and shared ID track
electron	muon	shared ID track
jet	electron	$\Delta R < 0.2$
electron	jet	$0.2 < \Delta R < 0.4$
jet	muon	NumTrack < 3 and (ghost-associated or $\Delta R < 0.2$)
muon	jet	$\Delta R <$ min(0.4, 0.04 + 10GeV/MuPt)
fat-jet	electron	$\Delta R < 1.0$

Table 5.4: Conditions of the overlap removal. These cuts are applied sequentially. ΔR is calculated using rapidity by default trough the formula $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\Phi)^2}$.

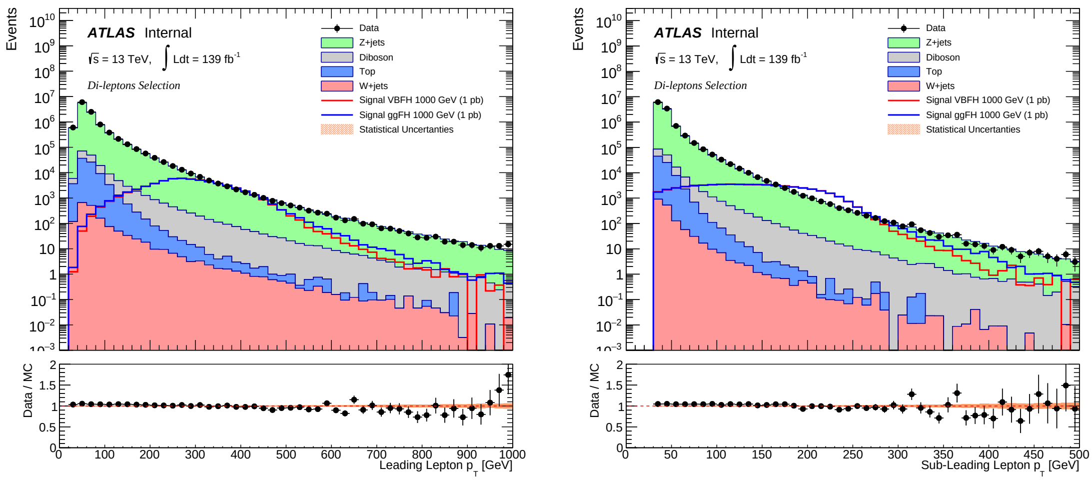
5.4 Event Selection

Topological cuts on the reconstructed objects are applied in order to reduce as much as possible the SM background and improve the sensitivity of the analysis.

The data used for the analysis have been collected during the full Run2 data taking period of the ATLAS experiment using single lepton or E_T^{miss} triggers, summarised in appendix A.1. Detailed studies on the triggers in the two leptons channel have been performed. A comparison between the different data period was done A using the data in order to study the efficiency of the different kind of triggers used. Part of my work has been also involved in the study of electron trigger at high-p_T range to prevent inefficiency of new triggers algorithms introduced during the 2018 period; details have been published [122] and summarised in Appendix B.

5.4.1 Di-leptons selection

In the 2 leptons channel the selection begins with the request of a pair of same flavour leptons ($l = e, \mu$). The pair is required to satisfy the *Loose* criteria (table 5.3) and to be isolated in the detector. Each lepton of the pair is required to have $p_T > 30 GeV$ to reduce fake leptons contribution (coming from the W+jets SM background) while preserving high efficiency for all signal samples (Figure 5.1).



(a) *Leading Lepton* (b) *Sub-Leading Lepton*

Figure 5.1: Data and MC comparison for the p_T of the 2 leptons selected.

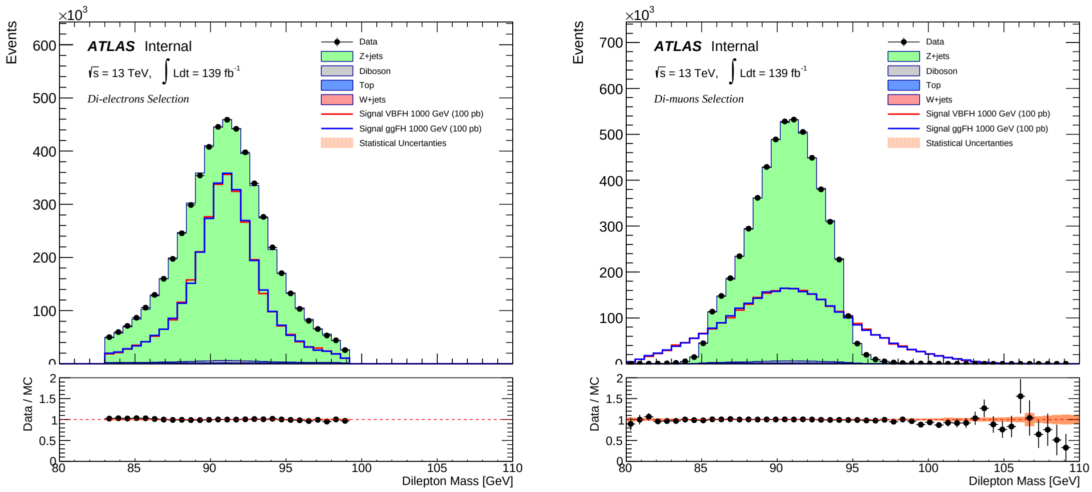
In the case of 2 muons they are required to have also opposite charges, while this is not applied in the case of 2 electrons since the reconstructed electrons have an higher probability of misidentification of the charge.

Then, the invariant mass of the di-leptons system is required to be around the Z boson peak, with a mass window cut of:

- ee pair $\longrightarrow 83 < m_{ee} < 99\ GeV$
- $\mu\mu$ pair $\longrightarrow (85.6 - 0.0117 \cdot p_T^{\mu\mu}) < m_{\mu\mu} < (94.0 + 0.0185 \cdot p_T^{\mu\mu})\ GeV$

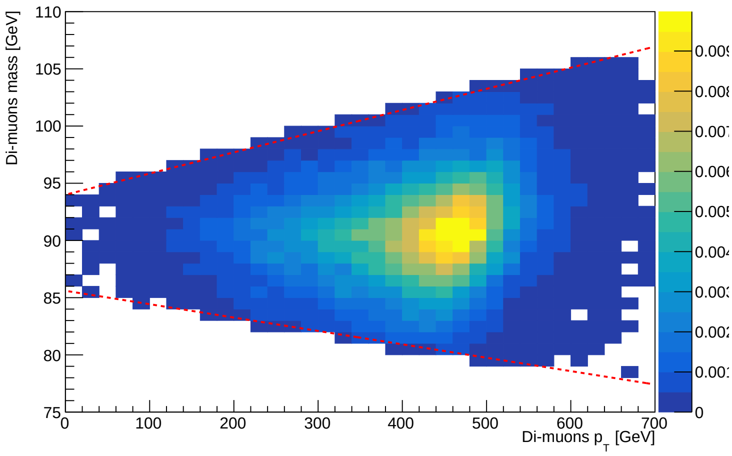
The distributions of the dilepton invariant mass is showed in Figure 5.2.

The p_T-dependent criteria applied in the di-muon channel is used to take into account the degradation of the mass resolution at high p_T range (Figure 5.3); both the windows are chosen to have $\sim 95\%$ of the signal efficiency as reported in the previous publication [53].



(a) *Di-electrons system* (b) *Di-muons system*

Figure 5.2: Data and MC comparison for the invariant mass of the di-electrons system (a) and of the di-muon system (b); the SM background has been normalised to the data and the signals have been scaled to a cross section of 100 pb; the di-muon spectra for the signals has a worst resolution since the degradation of the resolution at high-p_T.



(a)

Figure 5.3: Dimuon mass VS dimuon p_T after the mass window cut.

5.4.2 Leptons channel with Missing Energy

The 0 and 1 lepton channel are characterized by the presence of missing energy (E_T^{miss}) in the reconstructed final state. The 0 leptons channel events are selected requiring 0 loose leptons and $E_T^{miss} > 250 GeV$ due to high E_T^{miss} trigger thresholds, A.

For the 1 leptons channel, events must have exactly one tight electron or medium muon and a veto is applied on additional leptons. Moreover, $E_T^{miss} > 60\ GeV$ and the transverse momentum of the lepton and E_T^{miss} system (i.e. the reconstructed V_l), $p_T^{V_l} > 75\ GeV$ are applied.

Additional requirements on the kinematics of the E_T^{miss} respect to the other objects in the event are applied to further reduce the multijet background, details are summarised in table 5.5.

Event selection	0-leptons ($ZV \to \nu\nu V_h$)	1-leptons ($WV \to \ell\nu V_h$)
V_l selection	No *Loose* lepton $E_T^{miss} > 250\ GeV$ $p_T^{miss} > 50\ GeV$	1 *Tight* lepton $E_T^{miss} > 60\ GeV$ $p_T^{V_l} > 75\ GeV$
anti-QCD cuts	$\Delta\Phi(\vec{E}_T^{miss}, \vec{p}_T^{miss}) < 1$ $min\Delta\Phi(\vec{E}_T^{miss}, j) > 0.4$	$\Delta\Phi(\vec{E}_T^{miss}, l) < 1.5$ $\Delta\Phi(j_1, j_2) < 1.5$ $\Delta\Phi(l, j_{1/2}) > 1$ $\Delta\Phi(\vec{E}_T^{miss}, j_{1/2}) > 1$

Table 5.5: Overview of the $X \to VV \to V_l V_h$ selection criteria, see text for more details.

5.5 VBF/ggF classification

The VBF/ggF events classification has been exploited starting from the previous $36.1 fb^{-1}$ paper [53]. The VBF topology is characterised by two extra jets in the final states coming from the production mechanism (Figure 1.7c). The informations of the two jets can be used to identify the topology; indeed, they are produced with high energy and in the forward region of the detector, the jets pair has high invariant mass and high angular separation, as described in Section 1.6. A simple cut based approach was used to classify VBF and ggF/DY events [123, 124, 53].

The Figure 5.4 shows the efficiencies and the misidentification rate of this selection as function of the mass of the spin-0 resonance.

The cut based approach is mostly limited by:

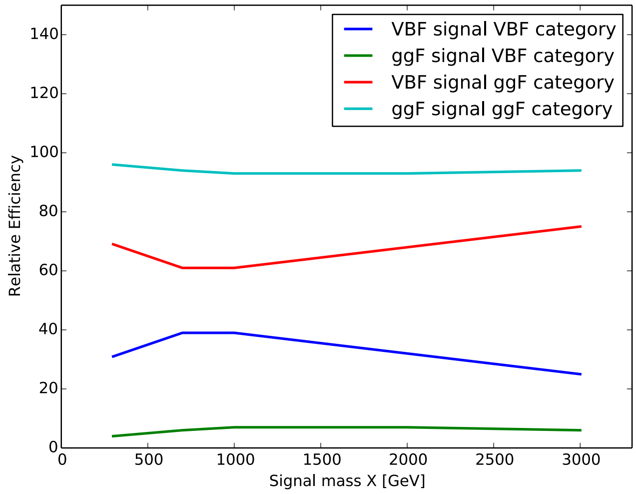
- geometric acceptance of the ATLAS detector, that allows to reconstruct both the VBF jets in 80% of events for low signal mass and in < 60% for 3 TeV signal mass;
- high mis-tagging of the truth VBF jets (VBF signal efficiency is < 40%).

5.5.1 VBF/ggF classification with a Neural Network approach

One of the main goals of my research activity has been the development of a new machine learning approach in the VBF/ggF event classification.

The VBF/ggF event classification can be seen as a binary classification problem in which 2 classes (VBF and ggF) has to be recognised using some features that characterize the two topologies. Usually, a Confusion Matrix (CM) is defined to quantify the goodness of the method used for the classification problem. The CM obtained in the case of the old cut-based classification is reported in table 5.6.

The aim of this new approach is mainly to increase the CM to increase the goodness of the classification. In the following, first the procedure used in the choice of the input variables and then the development of the best architecture will be discussed.



(a)

Figure 5.4: Signal efficiency and mis-tag rate in the VBF and ggF categories for the procedure described in [123].

	VBF class	ggF class
$VBF\ H$ 1000 GeV	38%	62%
$ggF\ H$ 1000 GeV	6%	94%

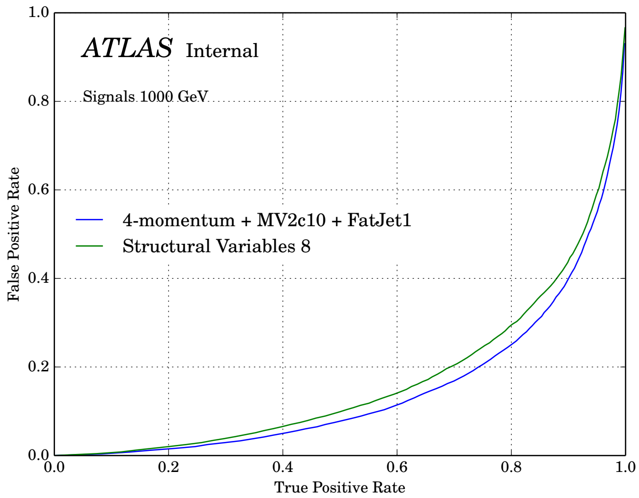
Table 5.6: Confusion matrix related to the cut-based categorisation for a 1000 GeV spin-0 resonance.

In high-energy physics, we can define:

- *Low-level variables*: simple and basic variables at reconstruction level, such as the components of the 4-momentum of the reconstructed objects;
- *High-level variables*: physics quantities that can be constructed starting from the low-level variables.

As discussed in recent work a deep learning architecture trained on a set of low-level variables can outperforms high level based classifiers. For this reason in order to choose the more appropriate set of input variables, we tested the discrimination performances of low-level variables (4-momentum of jets) and the high-level variables using a DNN architecture. High-level variables related to the two tag jets have been built and others related to the two Z bosons and combining the tag jets and the bosons (detailed informations are reported in appendix D).

A comparison of the DNN performances using the 4-momentum of the jets and a large set of high-level variables is showed in Figure 5.5. We trained the net over a wide range of high-level input variables and the performances has been always better using only low-level variables; the gain we have using low-level variables is reported on table **??** for different choice of working point.



(a)

Figure 5.5: Comparison of the DNN performances using the 4-momentum VS the Structural Variables approaches.

	False Positive Rate		
True Positive Rate	8 High level Variables	6jets 4-momentum	False Positive decrease
40%	6.7%	4.8%	40%
60%	14%	12%	17%
80%	30%	25%	20%

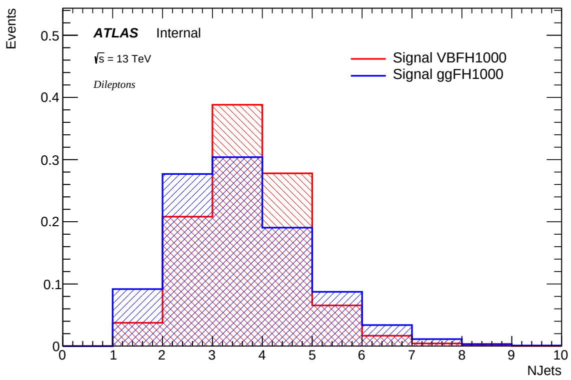
Table 5.7: False positive rate for different True Positive Rate WP with High-level and low-level variables; the last column represent the gain in terms of the False Positive decrease.

5.6 VBF/ggF classification: Recurrent Neural Network

After the preliminary test described in Section 5.5.1, we started to implement a ML based VBF/ggF classification using all the small-R jets in the event without any tagging procedure of the additional jets. Further, if one of the two VBF jets is outside the ATLAS geometric acceptance (see Figure 1.15) we can recover signal efficiency selecting also events with only one VBF jet.

In order to reach this aim the network architecture had to be able to deal with variable length input set and for this porpouse one possible solution is the *Recurrent Neural Network* (RNN) [103], introduced in Section 4.2.2.

We use the 4-momentum of the jets, p_T, η, ϕ, E. According to the intrinsic variable length of the input set (see Figure 5.6) the 4 input variables for each jet are passed to the RNN in a recurrent way.



(a)

Figure 5.6: Jets multiplicity of the two signal samples used for training the network.

To check the performances a RNN model is trained to separate the 1 TeV spin-0 VBF signal from the 1 TeV spin-0 ggF signal. The training has been made in 2-lepton channel. In the training and test phases all the events after the usual dileptons requirements have been considered.

5.6.1 Input set

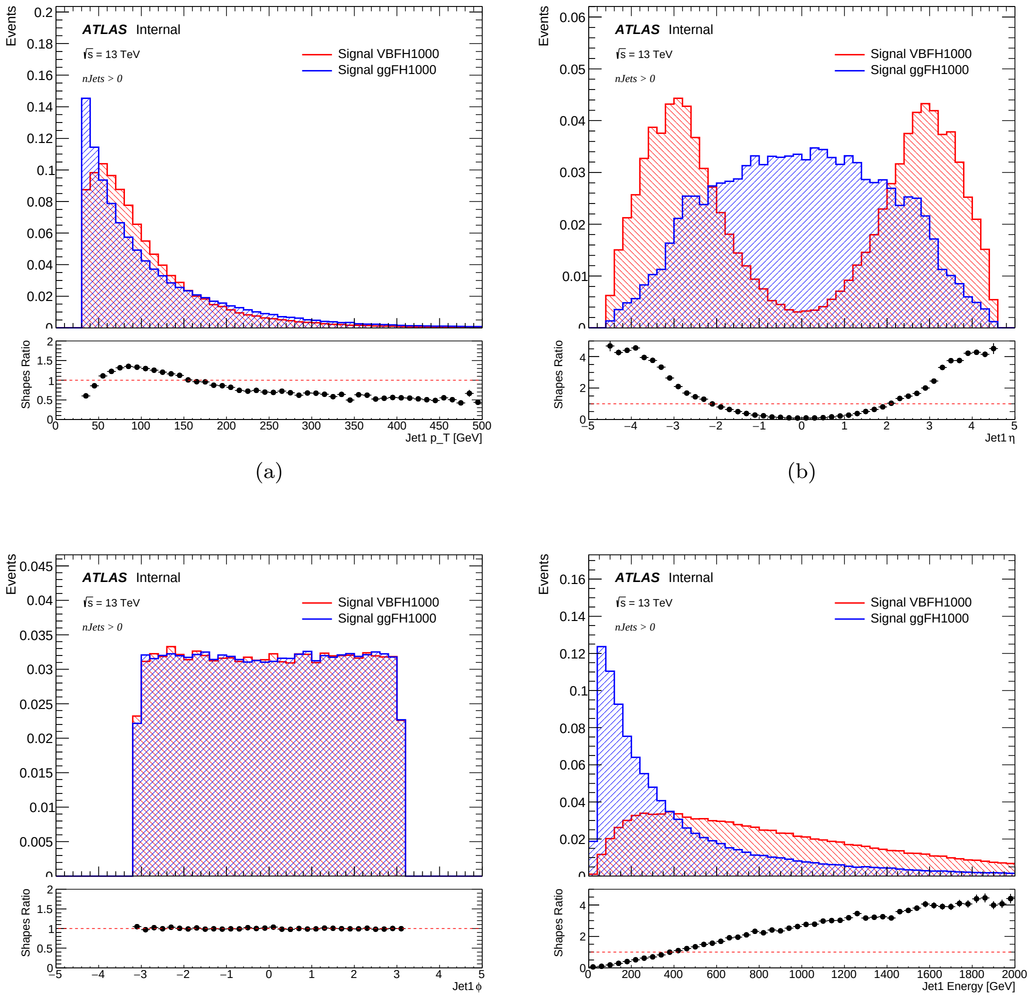
The input jets are chosen according to a loose pre-selection 5.3. For inputs jets to RNN, in order to avoid contamination of the signal jets coming from the V decay and to maximise the learning process of the network an overlap removal is applied:

- if there is a large-R jet in the event: small-R jets are discarded if they are matched to the leading large-R jet by $\Delta R < 1.0$.
- if there is not any large-R jet in the event: the leading two small-R jets in $|\eta| < 2.5$ are excluded from the list.

Then remaining small-R jets in $|\eta| < 4.5$ are considered as RNN inputs. The overlap removal procedure has been applied in order to simplify the inputs but the network still work with the same performances without this procedure.

The maximum number of input jets is fixed to 2 (i.e. we only use the leading two small-R jets after the overlap removal described above) in order to avoid the poor modeling of number of jets distribution. Moreover, the performances of the RNN show a negligible improvement if we use more jets as input E.3.

The input variables used in the training are shown in Figure 5.7 for the leading jet (Jet1) and in Figure 5.8 for the sub-leading jet (Jet2) if it exists in the event. The VBF-jet topology is quite clear in the input variables as shown in Figures 5.7-5.8. We expect to have VBF jets in the forward region of the detector with higher energy with respect to the usual extra jets we have in the other physics processes.



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

Figure 5.7: 4-momentum of the Jet1 used as RNN input for VBF (red) and ggF (blue) Higgs signal in 2-lepton channel.

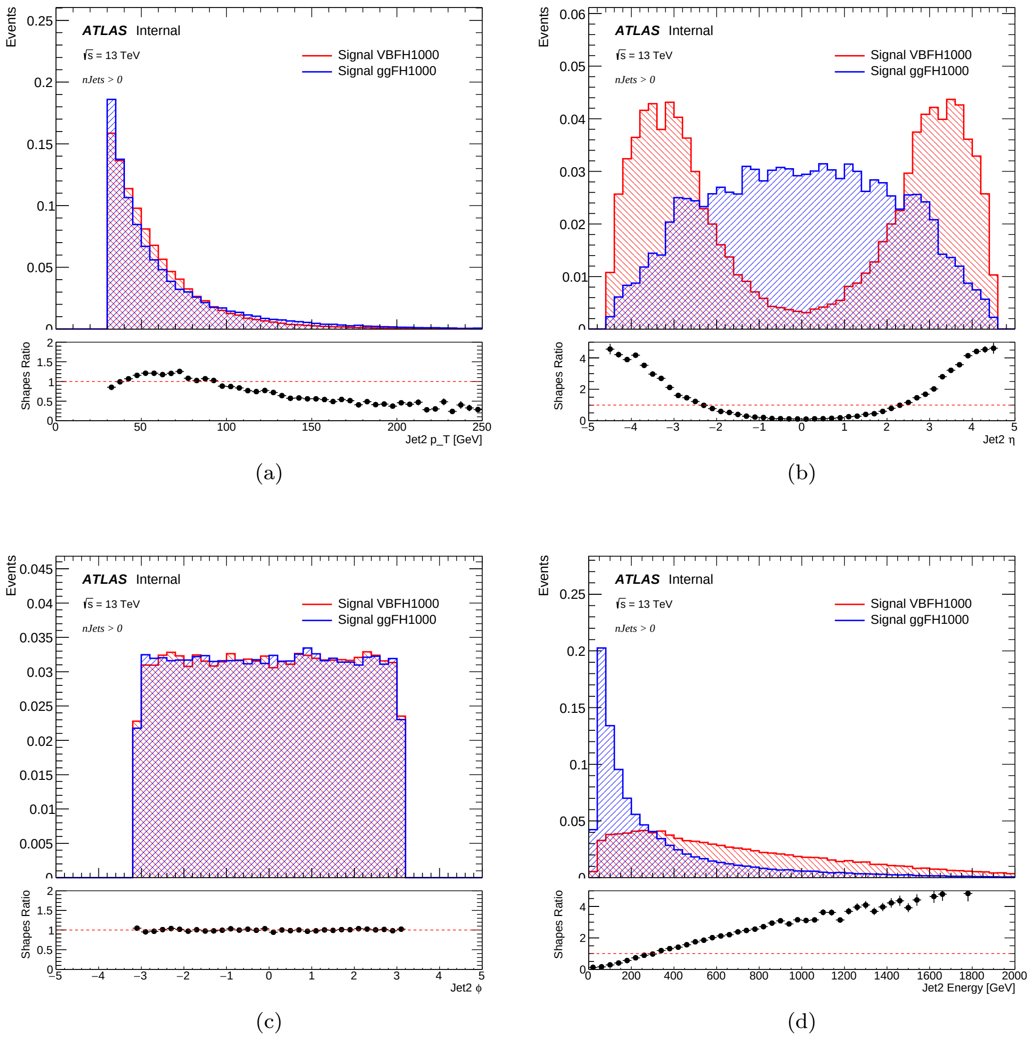
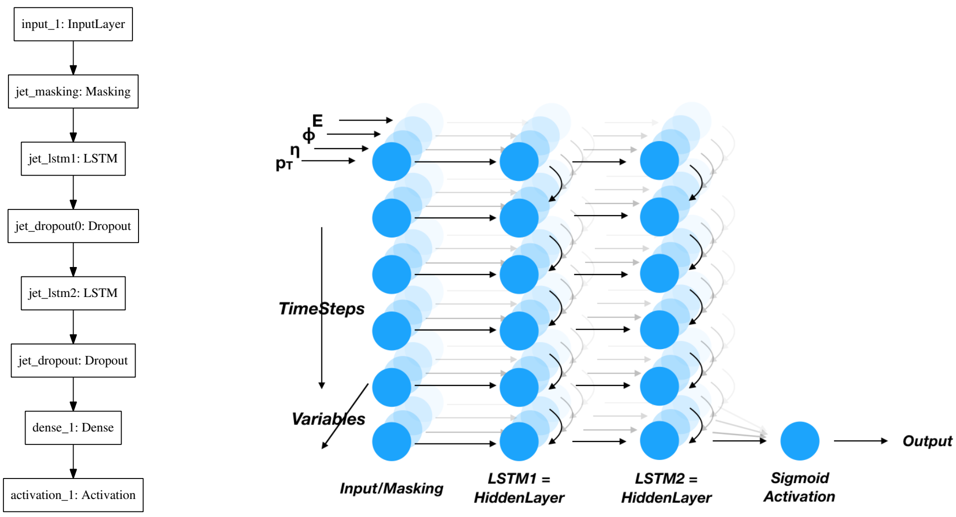


Figure 5.8: 4-momentum of the Jet2 used as RNN input for VBF (red) and ggF (blue) Higgs signal in 2-lepton channel.

5.6.2 RNN Architecture

A scheme of the architecture is reported in Figure 5.9 and some technical details are in Table 5.8. All the technical details are exploited in the dedicated Section 4.



(a) (b)

Figure 5.9: RNN architecture scheme (a) and pictorial representation (b).

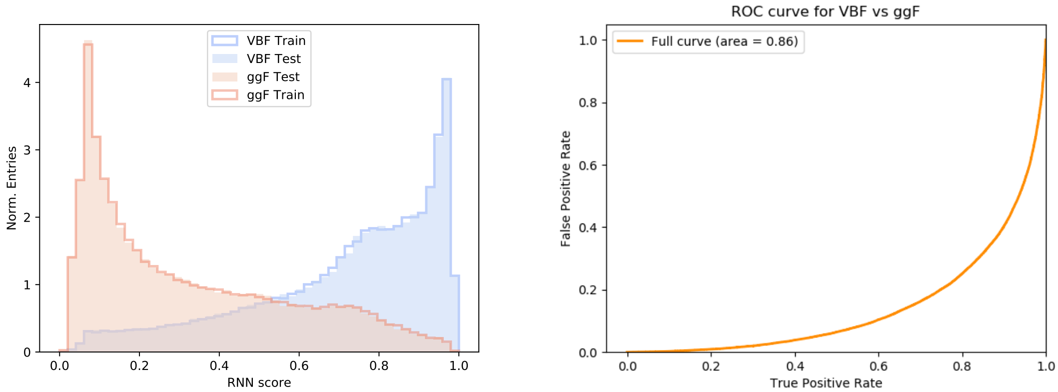
Loss function	binary crossentropy
Train:Test:Validation set	56 : 30 : 14%
Epochs	200
EarlyStopping Callback	after 10 unchanged iterations
LSTM Layer Activation	*tanh*
Output Layer activation	Sigmoid
# LSTM hidden layers	2
# cells / layer	25
Dropout	0.3
Optimiser	Adam
Batch size	512

Table 5.8: Technical setup of the RNN architecture.

The architecture has been designed to have just one standard neuron in the last layers so that to have just one score as output of the network; the distribution of the score obtained is showed in Figure 5.10-a comparing the training and the test dataset used. Figure 5.10-b shows instead the performances of the network throwing the ROC curve (Receiving Operating Curve).

Here, true and false positive rates are defined as:

$$TruePositiveRate = \frac{\#\ of\ truth\ VBF\ events\ with\ RNN > X}{\#\ of\ truth\ VBF\ events} \tag{5.1}$$



(a) (b)

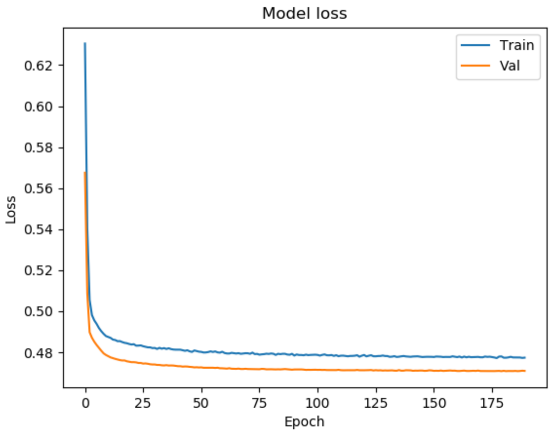
Figure 5.10: RNN score (a) and ROC curve performances.

$$FalsePositiveRate = \frac{\#\ of\ truth\ ggF\ events\ with\ RNN > X}{\#\ of\ truth\ ggF\ events}. \tag{5.2}$$

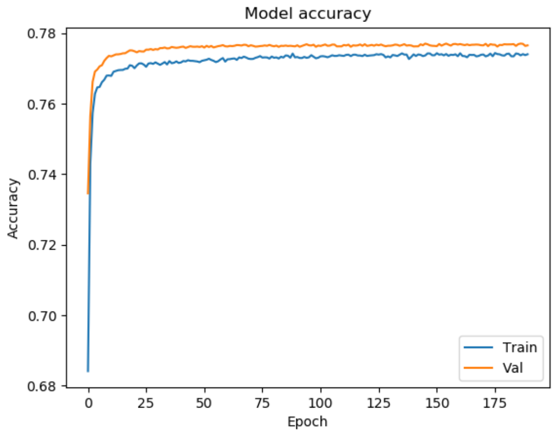
Figure 5.11 shows the stability and the goodness of the RNN model. Figures 5.11 show the loss function (a) and the accuracy (b) of the network as a function of the epochs of the training. The network does not show any sign of over-training since the validation loss is smaller than the training one and the accuracy is bigger than the training one during all the learning epochs.

If we split the whole dataset into training and test datasets in different random ways and we redo the training, the performances for each random case are found to be comparable (Figure 5.11-c); this shows that the network has no bias from the particular choice of the training set.

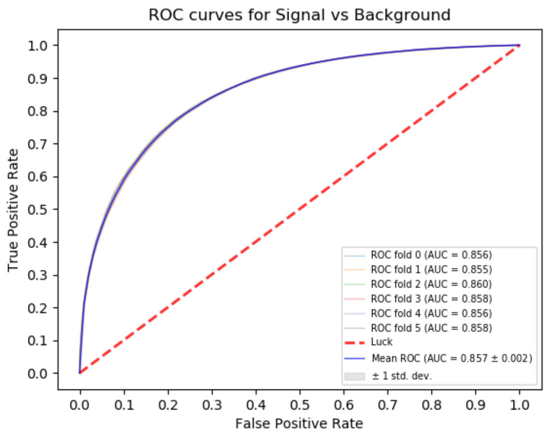
The usual *k-fold* validation test has been performed to check the dependency from the particular choice of the training/test datasets. We find the performances are consistent for 5 different choice of the training dataset, so no bias is introduced by the particular choice of the training dataset made.



(a)



(b)



(c)

Figure 5.11: Loss function (a) and accuracy (b) during the learning phase of the model; k-fold validation for the RNN model used (b) with signal = VBFH1000 and background = ggFH1000 (c).

5.7 RNN performances

5.7.1 Resonance mass dependency

The model trained on the 1 TeV signals is used in all mass ranges, since the VBF topology is expected to be independent on the signal mass. The 1 TeV hypothesis has been considered since the it is in an intermediate mass range between the resolved and merged regime of diboson system. Figure 5.12a shows the shapes of the RNN scores for the VBF/ggF signals with the different signal masses; the RNN model trained at 1 TeV is used. The separation power of the RNN score is quite good over the whole range of signals masses. The corresponding ROC curves for different mass points are shown in Figure 5.12b; the performances are even better as the resonance mass increases.

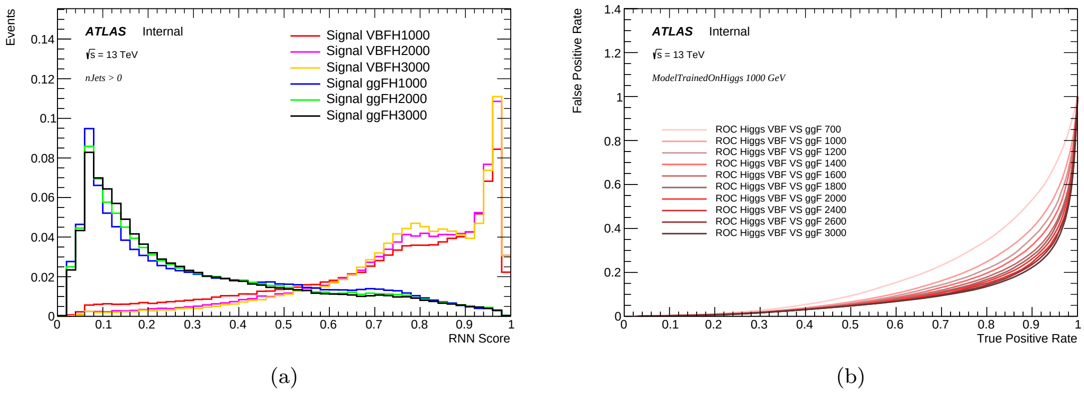
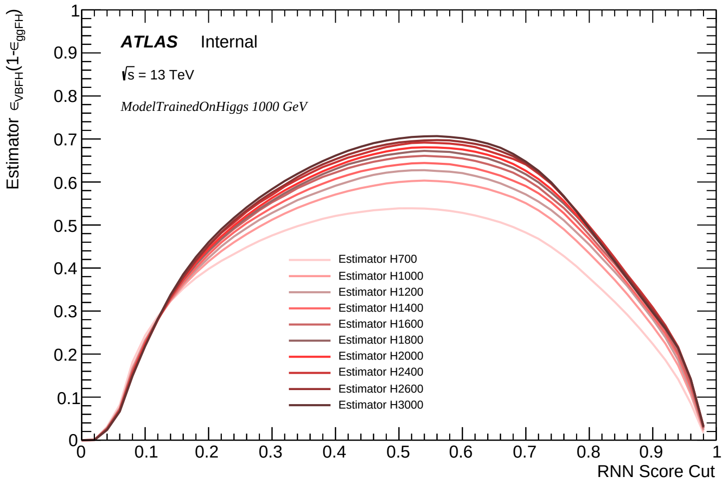


Figure 5.12: Shape distributions of the RNN Score for VBF and ggF signals with $mass = 1, 2, 3\ TeV$ (a) and corresponding ROC curves (b).

An estimator has been defined in order to quantify the separation power and the goodness of the CM obtained once a cut is fixed on the RNN score:

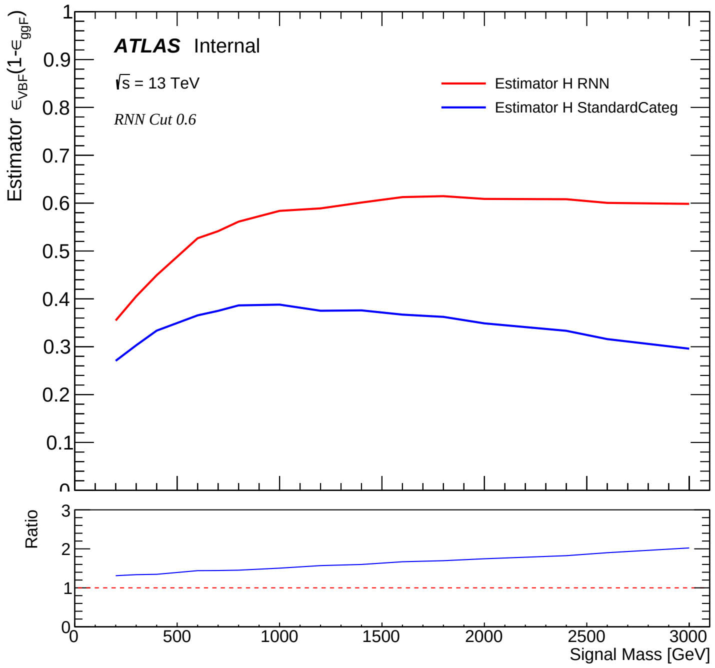
$$\sigma = \epsilon_{VBF} \cdot (1 - \epsilon_{ggF}) \tag{5.3}$$

where ϵ_{VBF} is the true positive rate and ϵ_{ggF} is the false positive rate. The Figure 5.12b shows the estimator varying the value of the cut on the score and for different mass points; the maximum of the estimator represents the best choice to optimise the CM and this is stable for the full mass range exploited ($\sim 0.6 - 0.7$). The Figure 5.14 shows the estimator with a fixed cut on the RNN score equal to 0.6 and varying the mass of the resonance.



(a)

Figure 5.13: Performances of the RNN for different signal masses represented in a estimator.

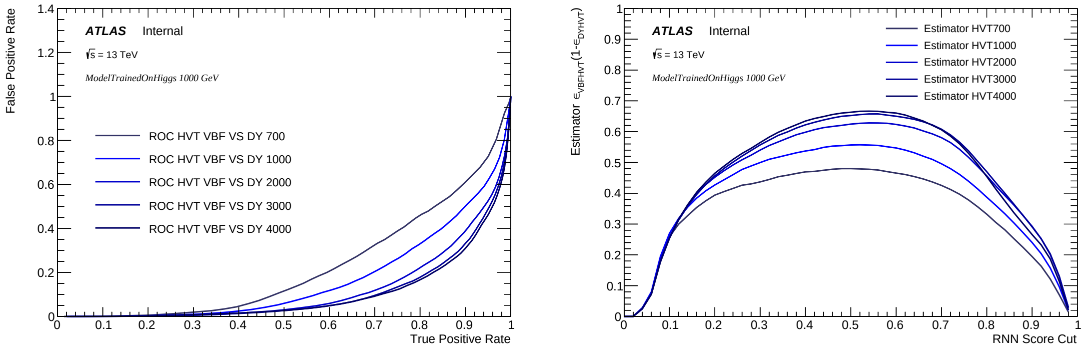


(a)

Figure 5.14: Performances of the RNN for different signal masses represented in a estimator for the two method, the RNN approach (red) and the cut based one (blue).

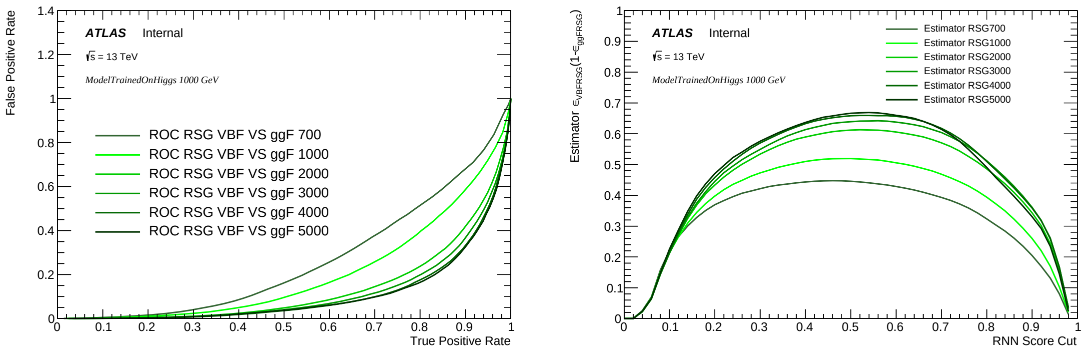
5.7.2 Resonance spin dependency

The RNN model trained on the 1 TeV Spin-0 signals has been used also for other spin hypotheses. Figures 5.15 and 5.16 show the ROC curves for spin-1 and spin-2 signals, respectively. The performances are good and similar to the Spin-0 case.



(a) (b)

Figure 5.15: Performances of the RNN for different Spin-1 signal masses represented in ROC curves (a) and estimator (b).



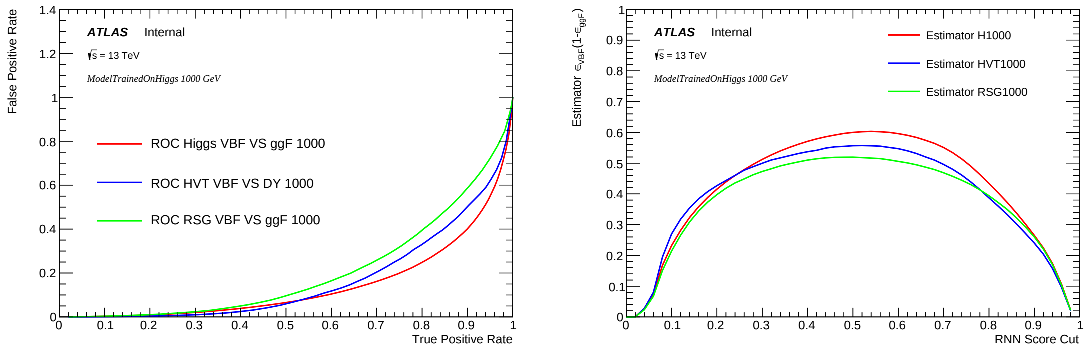
(a) (b)

Figure 5.16: Performances of the RNN for different Spin-2 signal masses represented in ROC curves (a) and estimator (b).

Figure 5.17 shows the comparison of the RNN performances for a benchmark mass of 1000 GeV and for the three different spin hypotheses. The Spin-0 is optimal since the training has been performed on that hypothesis; the VBF efficiency for the other spin hypotheses is 10% less respect to the Spin-0.

5.7.3 Application to 1- and 0-lepton channels

The RNN model trained on the 2-lepton signals is used also in the 0- and 1-lepton channels, as shown in Figure 5.18. The performances are almost identical to the 2-lepton channel as expected. This is expected since the VBF topology is completely independent from the number of leptons in the final state.

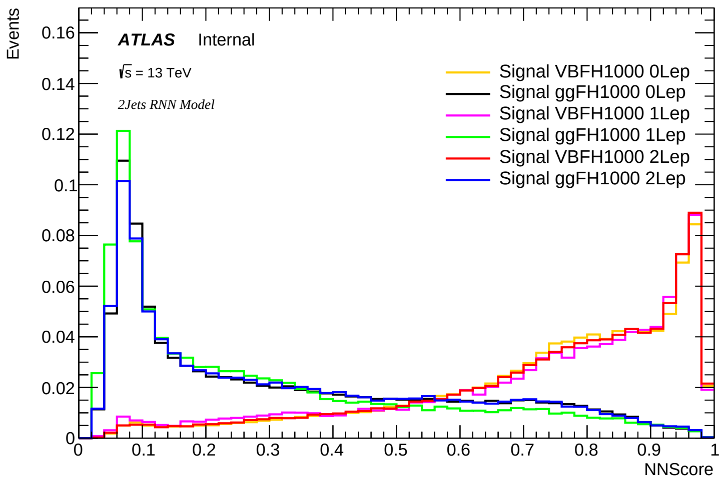


(a) (b)

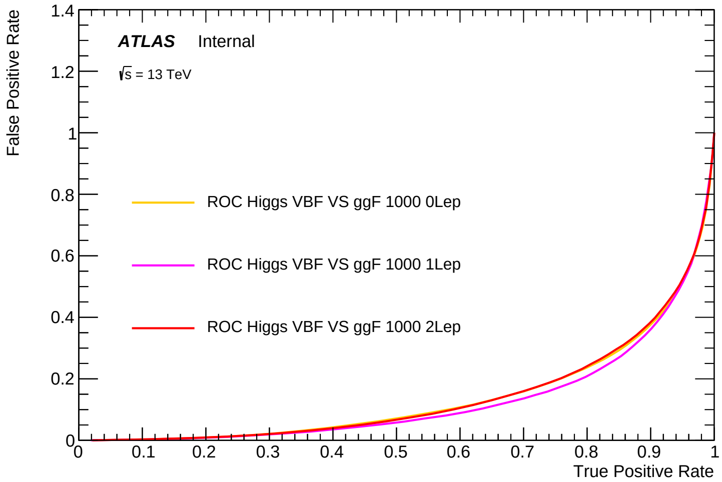
Figure 5.17: Performances of the RNN for different Spins signals for the 1000 GeV mass represented in ROC curves (a) and estimator (b).

Spin	VBF Efficiency			ggF/DY Contamination		
	RNN $>$ 0.6	RNN $>$ 0.7	RNN $>$ 0.8	RNN $>$ 0.6	RNN $>$ 0.7	RNN $>$ 0.8
Spin-0	73%	62%	46%	18%	11%	5.3%
Spin-1	64%	54%	40%	15%	8.1%	2.4 %
Spin-2	61%	53 %	42%	18%	11%	5.6%

Table 5.9: VBF efficiency and ggF/DY contamination for the benchmark mass hypothesis of 1000 GeV for the three different spins models.



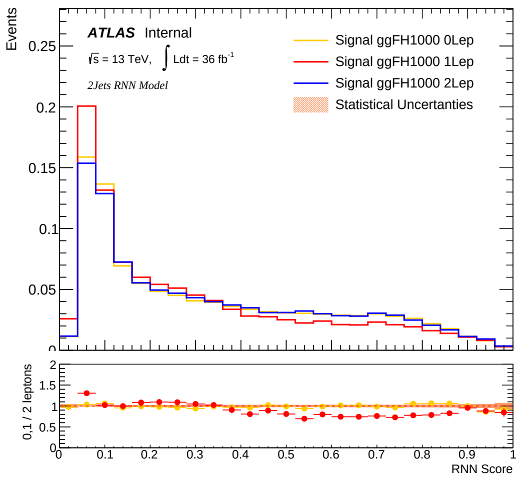
(a)



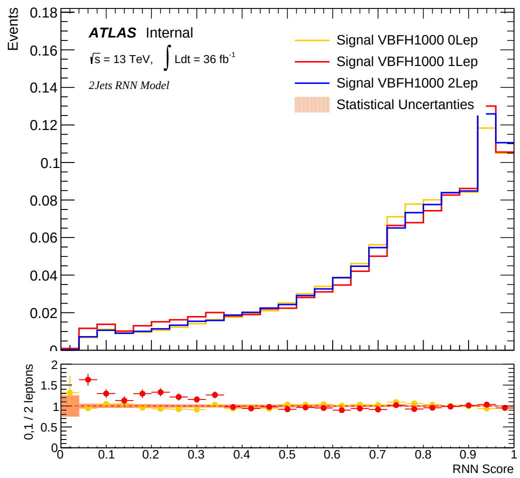
(b)

Figure 5.18: Performances of RNN model on the three 0/1/2 leptons channels.

In order to show better the behavior of the RNN score for the 3 different leptons channel the ratio of distributions of the scores has been evaluated and compared within the statistical uncertainties of the samples (5.19); a better consistency is observed for 0 and 2 leptons channel.



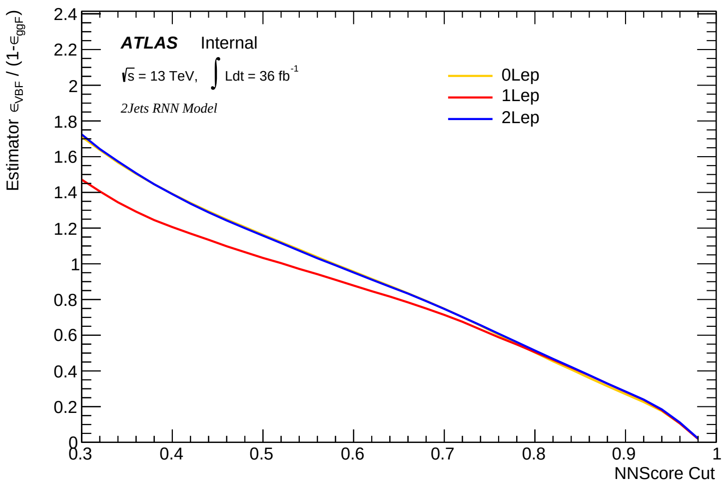
(a)



(b)

Figure 5.19: Comparison of the RNN scores for the 3 leptons channel; statistical error bands have been added in the ratio plot.

In order to quantify also the impact of this, the estimator given by the ratio of the VBF efficiency over the ggF rejection has been evaluated and compared for the 3 leptons channel; the estimator is consistent at the current WP of 0.8 (5.20).



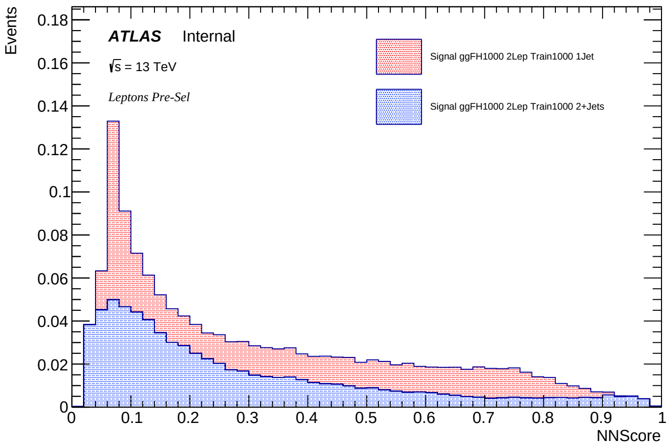
(a)

Figure 5.20: Comparison of the efficiency over rejection estimator.

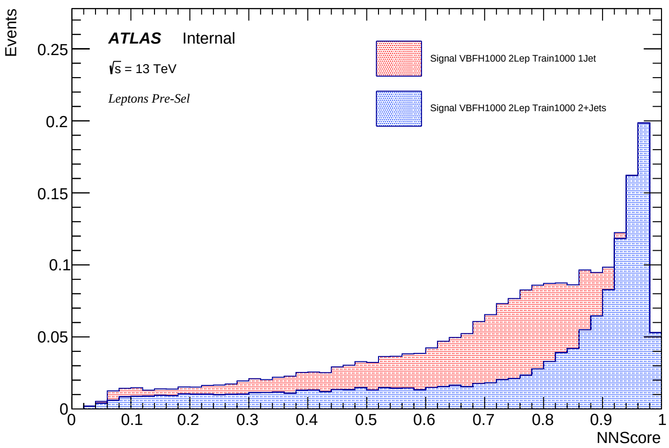
5.7.4 Events with 1 VBF jet

As pointed out in Section 5.5.1, one of the main improvement coming from the ML approach is the possible recovery of events with only 1 VBF jet.

The separate contributions of the $1Jet$ and $\geq 2Jets$ events to the RNN score shape is showed in Figure 5.21 for a 1 TeV signal sample. As expected, the network shows different behavior for these two events topology; the score of the VBF signal has, for instance, bigger values for the $\geq 2Jets$ events while smaller values for the $1Jet$ events.



(a)



(b)

Figure 5.21: RNN score for the $1-Jet$ and the $\geq 2Jets$ events.

5.7.5 Input features ranking

In this section the relative importance of the 4 input RNN jet variables (p_T, η, ϕ, E) has been investigated using the *permutation importance* method, implemented in the *Eli*5 library [125].

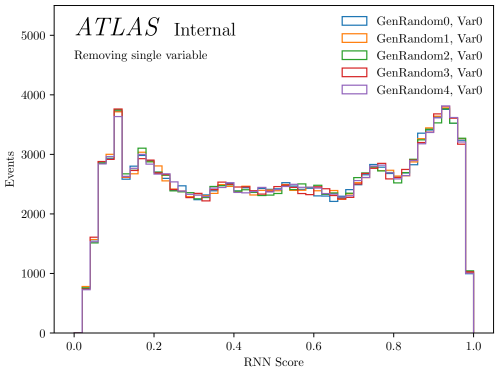
The idea is the following: feature importance can be measured by looking at how much the score (accuracy, loss, AUC, etc. decreases when a feature is not available.

To do that one can remove feature from the dataset, re-train the estimator and check the score. But it requires re-training an estimator for each feature, which can be computationally intensive. Also, it shows what may be important within a dataset, not what is important within a concrete trained model.

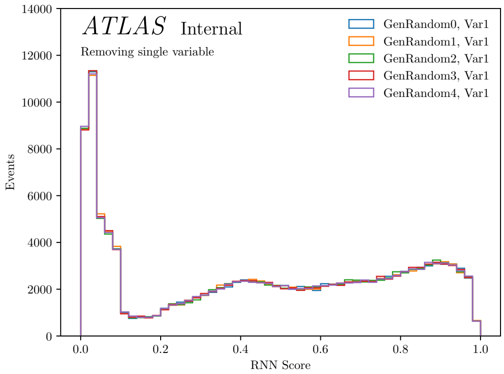
To avoid re-training the estimator we can remove a feature only from the test part of the dataset, and compute score without using this feature. It does not work as it is, because estimators expect feature to be present. So instead of removing a feature we can replace it with random noise - feature column is still there, but it no longer contains useful information. This method works if noise is drawn from the same distribution as original feature values (as otherwise estimator may fail). The simplest way to get such noise is to shuffle values for a feature, i.e. use other examples' feature values - this is how permutation importance is computed.

The method is most suitable for computing feature importances when a number of columns (features) is not huge.

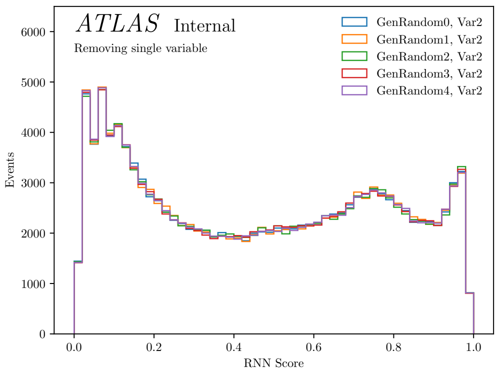
The 1 TeV Higgs signal events in the 2 leptons channel have been considered. The contribution of each of the 4 variables have been removed one by one by shuffling the events of the VBF and ggF signals. The random shuffling has been repeated 5 times (as the ELi5 library suggests). In Figure 5.22 the shapes of the RNN scores, evaluated for each random generation, are shown and good agreement between the random generation is observed.



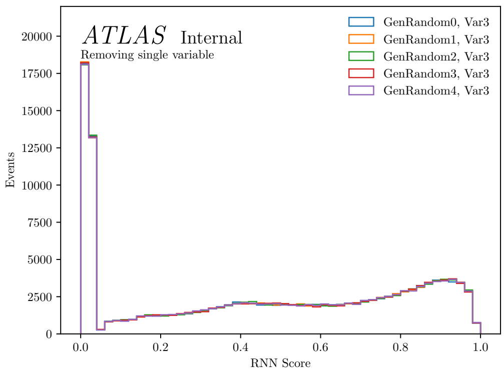
(a) p_T *removed*



(b) η *removed*



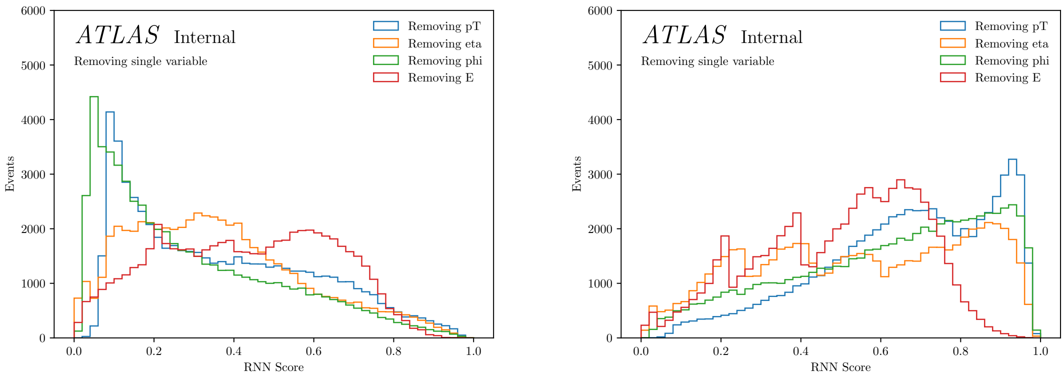
(c) ϕ *removed*



(d) *Energy removed*

Figure 5.22: RNN scores distribution comparison for the mixed VBF+ggF samples for 5 random generation of the shuffling of the 4 input variables.

For each of the 4 cases, in which the contribution of the single variable has been removed using the shuffling method, the RNN score has been built computing the mean of the 5 random generation. The distributions obtained are shown in figures 5.23 for the ggF and VBF signals.

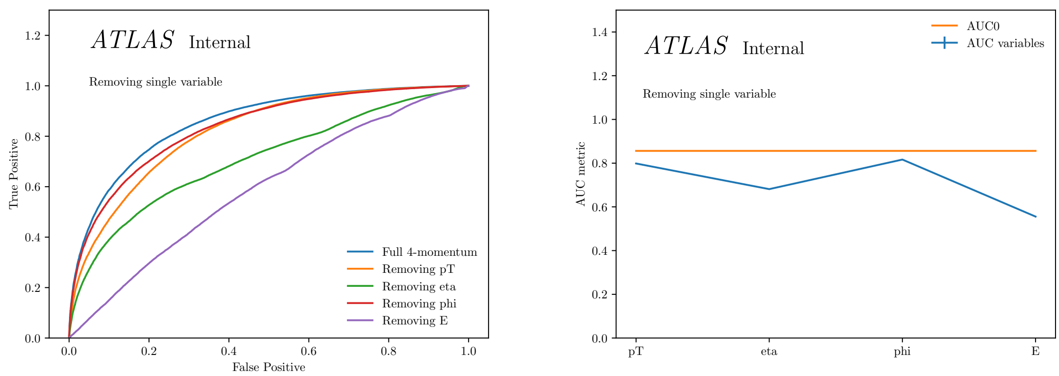


(a) (b)

Figure 5.23: RNN scores distribution comparison for each case in which one of the 4 input variable contribution has been removed for the ggF signal (a) and VBF signal (b).

It is quite evident that the RNN score shapes are effected less from the removing of the p_T and ϕ while the shape is strongly dependent from η and the energy, however none of the 4 variable has a negligible impact on the performances.

The ROC curves for each of the 4 cases have been built and compared with the ROC curve of the default model using the full 4-momentum; the AUC metrics (Area Under Curve) have been also compared, as shown in Fig. 5.24.



(a) (b)

Figure 5.24: ROC curves comparison (a) and AUC metrics comparison for the 4 cases with the default model.

A quantitative definition of feature importance is needed in order to give same features ranking. As we know, if a model performs worst with respect to another one the AUC is lower, so $1 - \mathrm{AUC}$ could be considered as an estimator of the error of the classification power of the model. The *feature importance* is defined as the factor of which the error on the classification increases if we remove that feature:

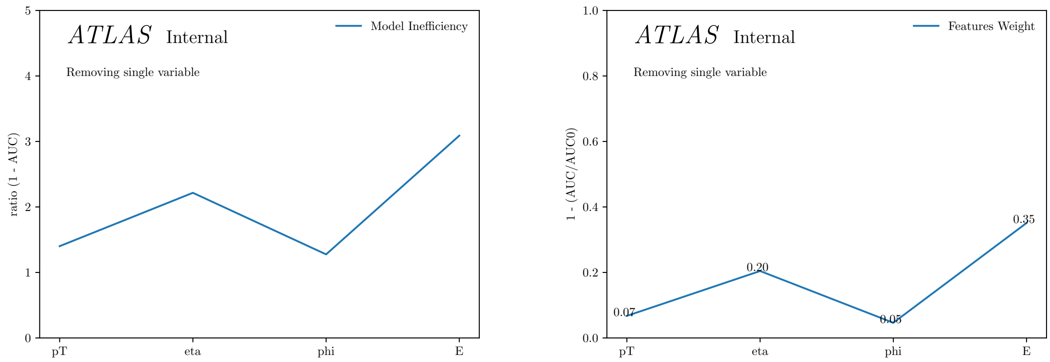
$$FeatureImportance = \frac{1 - \mathrm{AUC}}{1 - \mathrm{AUC}_{\mathrm{default}}} \tag{5.4}$$

an alternative definition useful for features ranking could be *feature weight* defined as:

$$FeatureWeight = 1 - \frac{\mathrm{AUC}}{\mathrm{AUC}_{\mathrm{default}}} \tag{5.5}$$

the feature importance and the feature weight for the 4 variables, p_T, η, ϕ, E are shown in figures 5.25. The ranking of the 4 input RNN variables is definitely:

$$energy \longrightarrow \eta \longrightarrow p_T \longrightarrow \phi \tag{5.6}$$



(a) (b)

Figure 5.25: Features importance (a) and features weight (b).

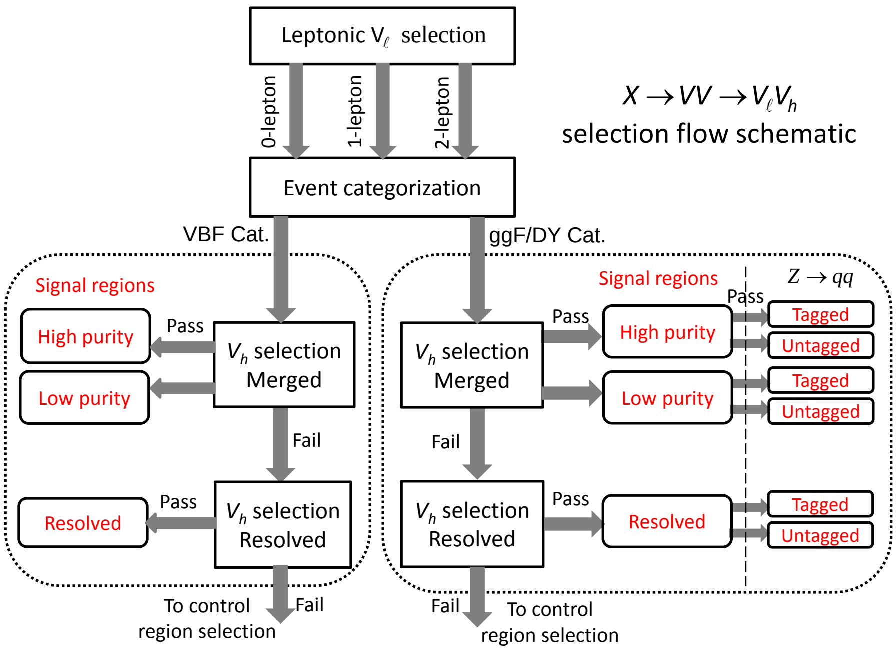
5.8 Analysis Strategy

The events collected are categorised according the number of Loose leptons (0, 1, 2) according the definition given in Section 5.3; this ensures the orthogonality of the channels.

Then, the events are categorised using the jets informations available in the events. First, the VBF/ggF(DY) mechanism production is exploited using a ML approach based on the RNN described in Section 5.6. After this categorisation the selection identifies jets coming from the hadronic decay of a boson V. Two techniques for the reconstruction of the signal jets are used:

- resolved topology: two small-R jets,
- merged topology: one large-R jet

the second technique is useful in the boosted regime in which the boson has high energy so that the decaying quarks have a high Lorentz boost and the are close each other . In this case, the decay production of the hadronisation of the two signal quarks are not completely reconstructed from a small-R jets and so a larger radius jet gives better performances in the reconstruction in terms of better mass and energy resolution on the reconstruction of the hadronic boson.



(a)

Figure 5.26: Analysis flow for SRs definitions.

5.8.1 Selection of the $W/Z \to jj$ candidates

For resonances mass below 1 TeV, the jets coming from the decay of a Z/W are still well separated and it is possible to identify them as two small-R jets. This resolved categories considers small-R jets with $|\eta| < 2.5$ and it is applied only for the 1 and 2 leptons channels.

5.8.2 $Z \to q\bar{q}$ selection

According to the SM branching ratios, the 21% of the cases of the decay of a Z boson contains b-quark while the fraction in the background is very small; this features is used in order to enhance the sensitivity using a b-jets category 3.3.4, defining:

- resolved tagged category: 2 signal b-jets are found, they are selected as the 2 signal jets
- resolved untagged category: less than 2 signal b-jets jets are found, the signal jets are selected as the 2 p_T-leading jets of the event.

5.8.3 $W \to qq'$ selection

In this case the resolved category is not splitted according the number of b-jets since a pair of b-jets is not expected this time. The 2 signal jets are selected as the 2 p_T-leading jets of the event.

5.8.4 Kinematic cuts

Once the pair of signal jets is identified, several kinematic cuts are applied in order to reduce the SM background. First, the p_T of the leading jet of the pair is required to be greater than 60 GeV and the sub-leading greater than 30 GeV. Than, a cut on the invariant mass of the di-jet system is applied in order to define the Signal Region (SR):

- $Z \to q\bar{q}$: $78 < m_{jj} < 105\ GeV$,
- $W \to qq'$: $68 < m_{jj} < 98\ GeV$,

A summary of the cuts applied is reported in table 5.10.

Selection		SR	Z CR (ZR)
$Z \to \ell\ell$	Number of Loose leptons	2	
	Same flavor	yes	
	Subleading lepton p_T	$> 30\,GeV$	
	dilepton invariant mass	$83 < m_{ee} < 99\ GeV$ $-0.01170ptll + 85.63 < m_{\mu\mu} < 0.01850ptll + 94.00gev$	
	Opposite sign	For $\mu\mu$ channel only	
Pass VBF selection		no (yes) for DY/ggF (VBF) category	
$W/Z \to jj$	Num of signal small-R jets	2	
	Leading jet pt	$> 60\,GeV$	
	Subleading jet pt	$> 30\ geV$	
	$Z \to q\bar{q}$	$78 < m_{jj} < 105\ GeV$	$60 < m_{jj} < 68\ GeV$ or
	$W \to q\bar{q}$	$68 < m_{jj} < 98\ GeV$	$105 < m_{jj} < 150\ GeV$
	Num. of b-tagged jets	For $Z \to jj$: ≤ 1 $(= 2)$ for untagged (tagged) category	
Topology cut	$\min\left(p_{T,ll}, p_{T,J}\right)/m_{llJ}$	$> 0.35(0.25)$ for the DY/ggF (VBF) category	

Table 5.10: Event selection summary for resolved analysis in 2-lepton channel.

5.8.5 Selection of the $W/Z \to J$ candidates

For high mass resonances ($\sim> 700GeV$), the decay of the hadronic boson could be often reconstructed as a unique large-R jet. So, also a "merged" category is considered in order to enhance the analysis sensitivity of the resolved category. For this aim, a large-R jet is required in the event with $p_T > 200\ GeV$ and $|\eta| < 2.$. When more than 1 large-R jet is reconstructed the p_T-leading one is selected as the signal jet.

The boson tagging procedure, based on the jet mass and jet substructure D_2 variables, is than applied on the signal candidate jet (Section 3.3.3). The boson tagger is a 2-variables based tagger (mass and D_2). The first one is used on order to define the SR, events around the W/Z mass window are in the signal region, while the D_2 is used in order to split the merged categories in 2 sub-categories according the purity of the signal:

- merged High Purity (HP) region: the event pass the D_2 cut,
- merged Low Purity (LP) region: the event does not pass the D_2 cut.

The boson tagger has been also optimised separately in order to identify $Z \to q\bar{q}$ and $W \to qq'$ candidates, so both the selections are considered. As for the resolved category, also the the b-jets multiplicity is used in order to maximise further the sensitivity. The b-jets multiplicity is counted using the *VR track jets* associated to the large-R jet 5.3. In particular:

- merged tagged category: 2 b VR track jets are found inside the signal large-R jet.
- merged untagged category: less than 2 b VR track jets are found inside the signal large-R jet.

A summary of the cuts applied is reported in table 5.11.

Selection		SR		ZCR (ZR)	
		HP	LP	HP	LP
$Z \to \ell\ell$	Number of Loose leptons	2			
	Same flavor	yes			
	Subleading lepton p_T	$> 30\,GeV$			
	dilepton invariant mass	$83 < m_{ee} < 99\ GeV$ $-0.01170 \cdot p_{T,ll} + 85.63 < m_{\mu\mu} < 0.01850 \cdot p_{T,ll} + 94.00\ GeV$			
	Opposite sign	For $\mu\mu$ channel only			
Pass VBF selection		no (yes) for the DY/ggF (VBF) category			
$W/Z \to J$	Num of large-R jets	≥ 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 $Z \to J$: ≤ 1 ($= 2$) for untagged (tagged) category			
Topology cut	$\min(p_{T,ll}, p_{T,J})/m_{llJ}$	$> 0.35(0.25)$ for the DY/ggF (VBF) category			

Table 5.11: Event selection summary for merged analysis in 2-lepton channel.

5.8.6 Signal Efficiencies in SRs

Signal selection efficiencies depend on the signal model, the production process and the mass of heavy resonances. Figures 5.27, 5.28 and 5.29 show the acceptance times efficiency ($A \times \epsilon$) of the signal events from MC simulations as a function of the resonance mass for (a) ggF/DY and (b) VBF production, combining all SRs of both ggF/DY and VBF categories of both resolved and merged analyses. The $A \times \epsilon$ curves are largely determined by the merged analyses. The resolved analyses contribute only at the low mass region, up to approximately 1 TeV.

Large differences in $A \times \epsilon$ shown in the figures for different resonances are the results of different spins of these resonances. The spin-0 radions are produced with isotropic angular distributions for both ggF and VBF production. On the contrary, the spin-1 HVT resonances and spin-2 RS gravitons are produced more centrally (more forward) for ggF/DY (VBF) production. These different angular distributions lead to very different efficiencies of the $R_{p_T/m}$ requirement. Moreover, the angular requirements between jets and E_T^{miss} in the 0-leptons channel are more efficient for DY production of HVT resonances than for ggF production of radion and graviton due to the difference in the color factors between the initial-state quarks and gluons.

Signal contributions from $W \to \tau\nu \to \ell\nu's$ decays are included in the 1-leptons channel, but not for the 2-leptons channels. Approximately 10–12% of the signal events in SRs are from $W \to \tau\nu \to \ell\nu's$ decays in the 0-leptons channel. These events have similar mass distributions as those from $W \to \ell\nu$ decays. In the 2-lepton channel, signal contributions from $Z \to \tau\tau \to \ell\ell\nu's$ decays are suppressed by the small $\tau\tau \to \ell\ell\nu's$ branching ratio and the Z-boson mass requirement. They are found to be negligible. The 0-leptons channel targeting the $X \to ZV \to \nu\nu qq$ signal should also be sensitive to the $X \to WV \to \ell\nu qq$, $\tau\nu qq$ signal due to either the inefficiency of the lepton veto or the lack of the τ veto. This additional "cross-channel" signal contribution is neglected in this search.

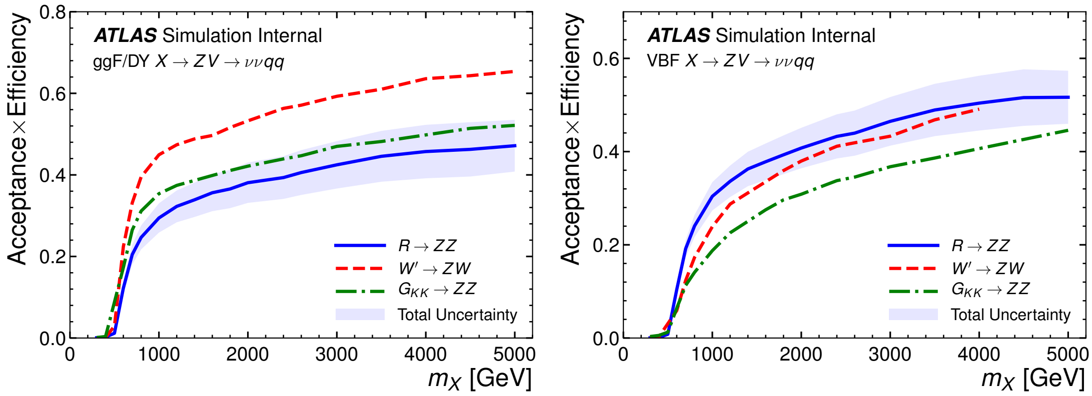


Figure 5.27: Selection acceptance times efficiency for the $X \to ZV \to \nu\nu qq$ signal events from MC simulations as a function of the resonance mass for (a) ggF/DY and (b) VBF production, combining HP and LP signal regions. The light blue band represents the total statistical and systematic uncertainties for the radion model, and the total uncertainties are similar for the other signal models.

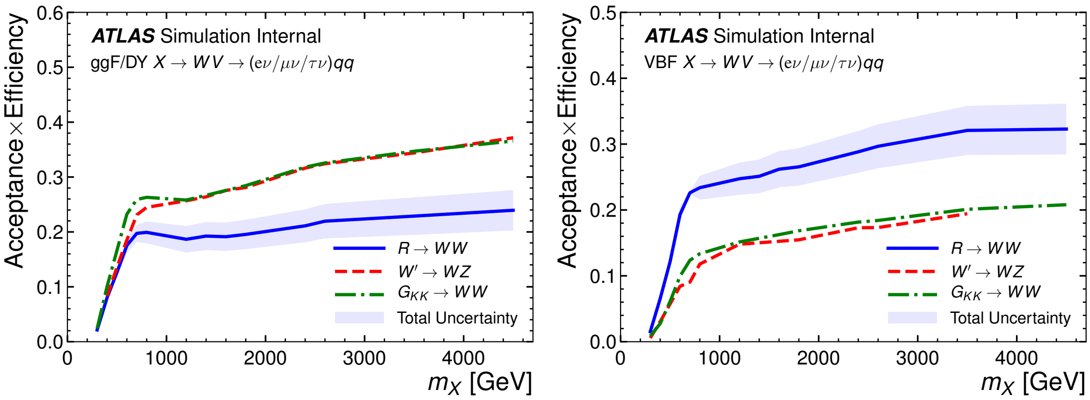


Figure 5.28: Selection acceptance times efficiency for the $X \to WV \to (e\nu/\mu\nu/\tau\nu)qq$ signal events from MC simulations as a function of the resonance mass for (a) ggF/DY and (b) VBF production, combining all SRs of both ggF/DY and VBF categories of both resolved and merged analyses. Signal contributions from $W \to \tau\nu$ decays are included in the efficiency calculation. The light blue band represents the total statistical and systematic uncertainties for the radion model, and the total uncertainties are similar for the other signal models. The "bump" structure around 800 GeV is due to the falling off of the resolved analysis.

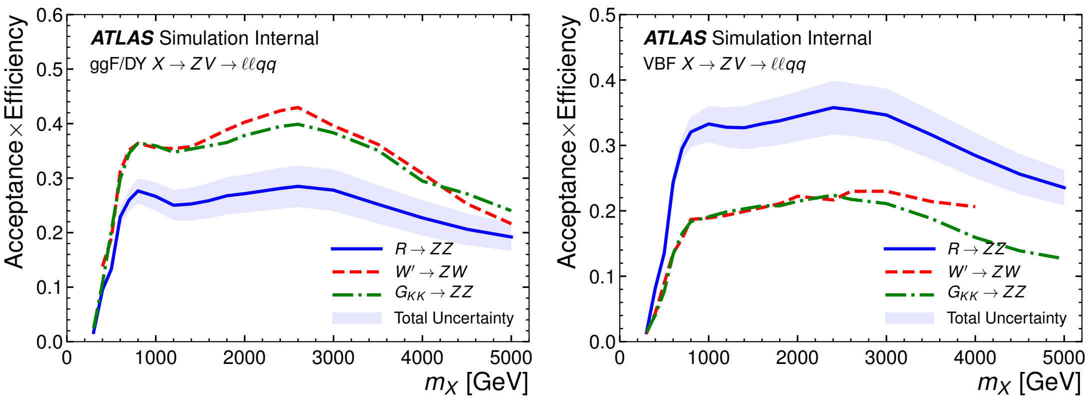


Figure 5.29: Selection acceptance times efficiency for the $X \to ZV \to \ell\ell qq$ signal events from MC simulations as a function of the resonance mass for (a) ggF/DY and (b) VBF production, combining all SRs of both ggF/DY and VBF categories of both resolved and merged analyses. The light blue band represents the total statistical and systematic uncertainties for the radion model, and the total uncertainties are similar for the other signal models. The decreases in efficiencies for resonance masses above approximately 2.5 TeV are due to the merging of electrons from the highly boosted $Z \to ee$ decays. The "bump" structure around 800 GeV is due to the falling off of the resolved analysis.

5.9 Control regions

The control regions are defined in a way similar as much as possible to the signal regions but requiring no signal or tiny signal contamination. Usually, a control region is defined applying the same sets of cuts of the related signal region but one which is orthogonal in order to make sure to not have much signal. An example of CR definition is given by reverting the mass window on the hadronic boson mass.

Once the CRs are defined they will be used in the statistical interpretation of the analysis in order to constraint the main background and improve the background description in the SRs.

In the 2 lepton channel, the main backgrounds are represented by the $V + jets$ and $t\bar{t}$ SM production as described in Section 1.4.

Dedicated CRs are defined for the $Z + jets$, $W + jets$ and $t\bar{t}$. Common CRs are shared between the 3 leptons channels, but independent CRs for merged and resolved regime are defined in order to take into account possible mis-modelling of the background depending on the p_T range of the jets.

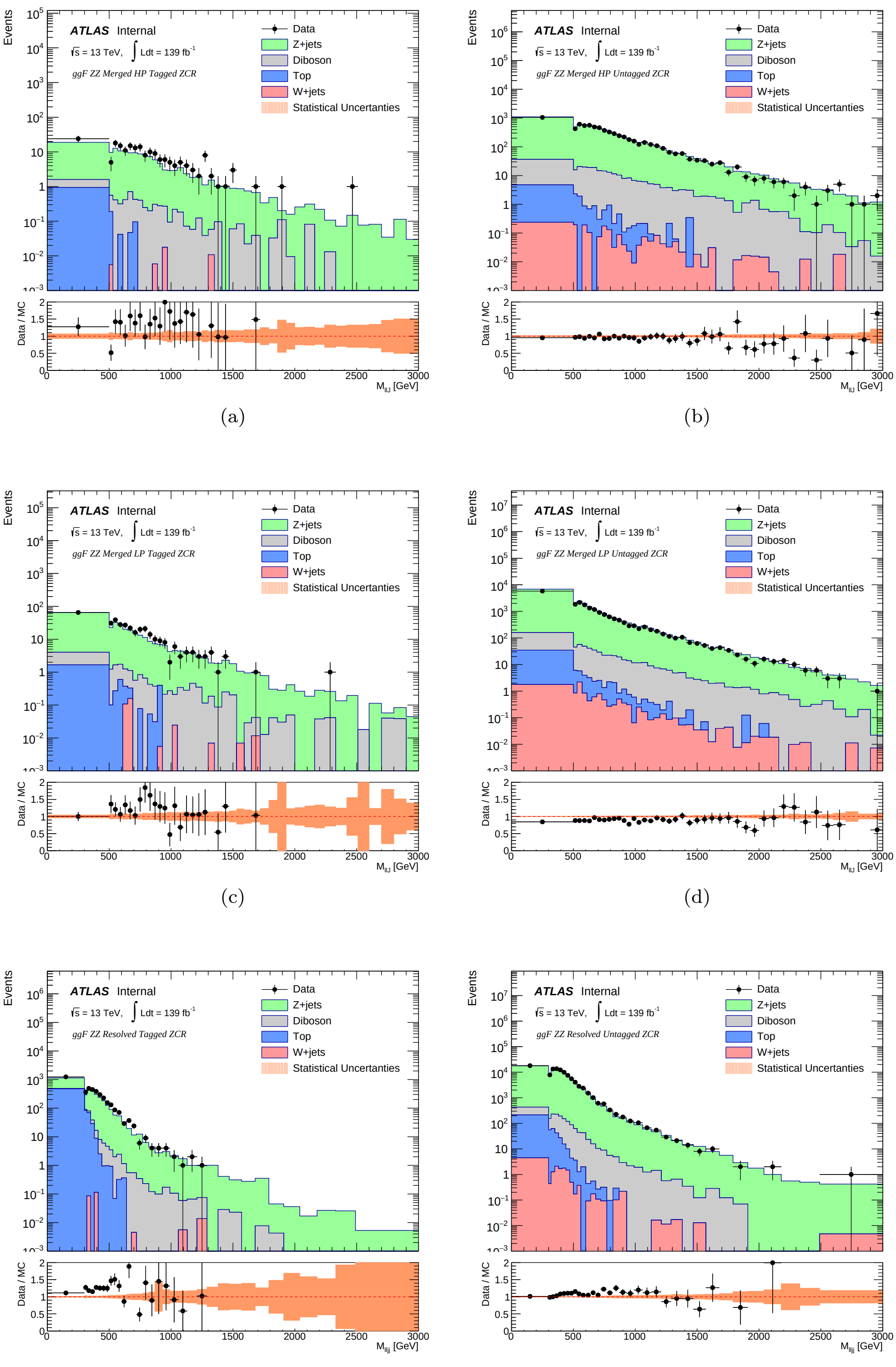
Z+jets (ZCR) and W+jets (WCR) control regions

The $Z + jets$ and the $W + jets$ control regions are defined using the same cuts of the SRs but using the mass side-band of the $Z/W \rightarrow qq$ system reconstruction; the ZCR is defined in the 2 leptons channel while the WCR in the 1 lepton. For the merged regime the side-bands of the boson tagger are used, so the range are p_T dependent, while for the resolved regime fixed side-bands are used, $50 < m_{jj} < 62||105 < m_{jj} < 150\ GeV$.

$t\bar{t}$ control region (TCR)

The top control region is defined only in the 1 lepton channel.

This region is defined requiring one extra b-jet respect the usual cuts of the SRs. An extra small-R jet passing the b-tagging criteria is required in the resolved regions while an extra VR track jet passing the b-tagging criteria is required in the merged regions.

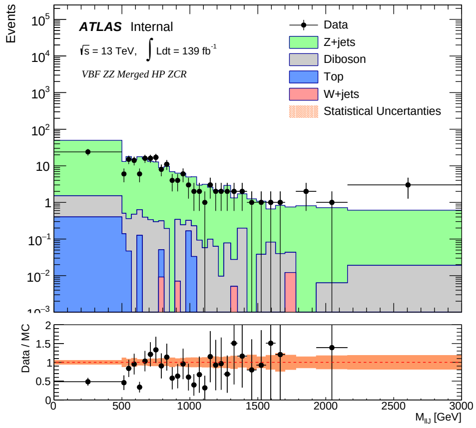


(a) (b)

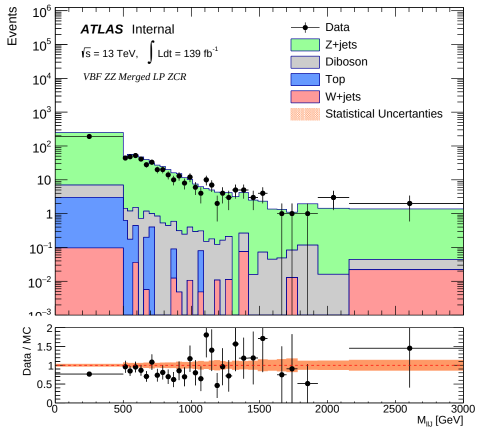
(c) (d)

(e) (f)

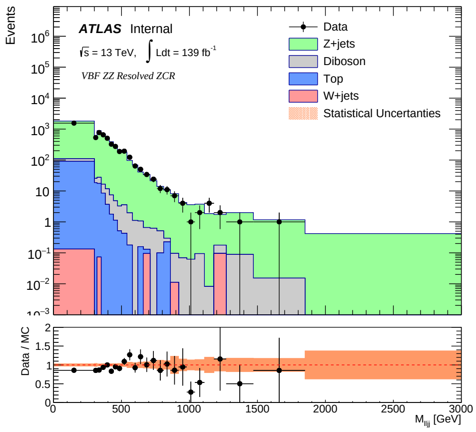
Figure 5.30: Data and MC comparison for the final discriminant in the ZCRs of the 2 leptons channel for the ggF selection.



(a)

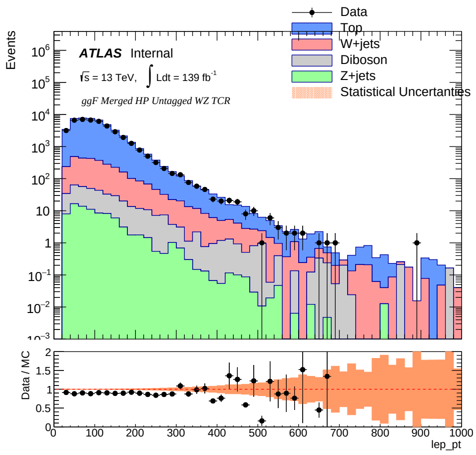


(b)

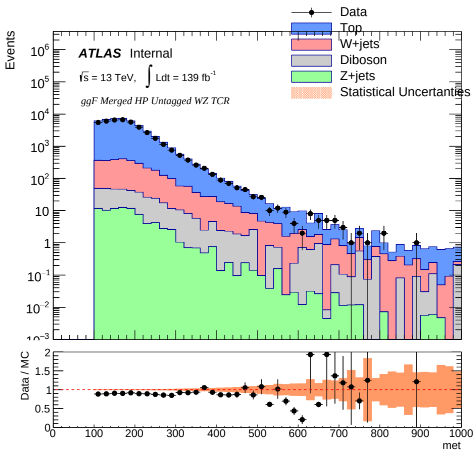


(c)

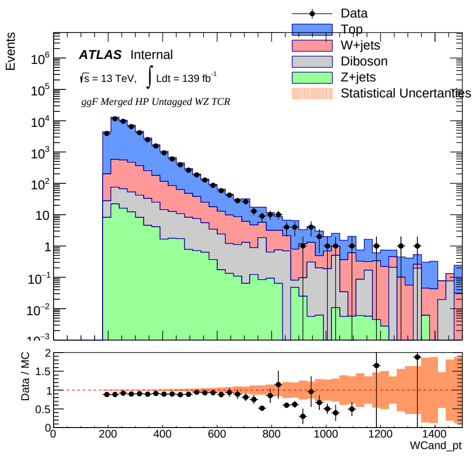
Figure 5.31: Data and MC comparison for the final discriminant in the ZCRs of the 2 leptons channel for the VBF selection.



(a)



(b)



(c)

Figure 5.32: Data and MC comparison for some kinematic variables in the TCRs of the 1 lepton channel for the ggF selection.

5.10 Systematic Uncertainties

The systematics uncertainties could affect the sensitivity of the search through the impact on the estimation of the background (MC based), the estimation of the signal and of the distribution of the final discriminant.

The sources of these uncertainties could be classified in two groups:

- experimental systematics, related to the detectors effect and reconstruction performances;
- theoretical systematics, coming from the limitation of the MC used for the simulation both for the background and the signals.

5.10.1 Experimental uncertainties

The experimental uncertainties arise from general informations of the events used as the luminosity measurement and the trigger efficiency in the selection of the events; they also are related to the identification and reconstruction of the physics objects used in the analysis, leptons, missing energy and jets.

The uncertainty in the combined 2015-2018 integrated luminosity is 1.7%. It is derived from the calibration of the luminosity scale using a methodology based on the x-y beam-separation scans, similar to that detailed in Ref. [126], and using the LUCID-2 detector for the baseline luminosity measurement [127]. A variation in the pile-up re-weighting of MC events has been considered to cover the uncertainty in the ratio of the predicted and measured inelastic cross section [128].

The efficiencies of the lepton triggers are high and their related uncertainties are negligible. The modelling of the electron and muon reconstruction, identification and isolation efficiencies are studied with a tag-and-probe method using $Z \to ll$ events in data and simulation [129]-[130]. Small corrections are applied to the simulations to better model the performance seen in data. These corrections have associated uncertainties on the order of 1%. Uncertainties in lepton energy (or momentum) scale and resolution, especially for muon momentum resolution (3%), are also taken into account.

Uncertainties for the jet energy scale and resolution for small-R jets are measured using MC simulation and in-situ techniques [131]. For central jets, the total relative uncertainty on the jet energy scale varies in the range $1-4\%$ for $p_T > 20\ GeV$. The uncertainty on the jet energy resolution ranges from 20% for jets with a p_T of 20 GeV to less than 5% for jets with $p_T > 200\ GeV$.

Uncertainties on the scale of large-R jet p_T is estimated by comparing the calorimeter and track-based energy and mass measurements in data and simulation [132]. The precision of the relative jet energy scale is $1-2\%$ and of $2-10\%$ for the mass scale for $200\ GeV < p_T < 2\ TeV$. The jet energy resolution uncertainty is estimated by degrading the nominal resolution by a 2%.

The efficiency of W/Z-tagging based on cuts on jet mass and $D^{\beta=1}$ is estimated in data using control sample and corrected by comparing it with simulation. The efficiency to W/Z-induced jet signal is estimated by $t\bar{t}$ control sample, while the efficiency to singlel q/g background is estimated by di-jet sample. The effects of experimental and theoretical uncertainties on the efficiency scale factor has been taken into account.

Uncertainties in the efficiency for tagging b-jets and in the rejection factor for light jets are determined from $t\bar{t}$ control samples [133]-[134]. The total uncertainties are $1-10\%$, $15-50\%$, and $50-100\%$ for b-jets, c-jets, and light-flavour jets respectively. The b-tagging efficiency uncertainty, which mainly affects the tagged signal regions, has an effect of about 10%.

An uncertainty for the E_T^{miss} trigger has been considered, this arises from the estimation of the difference between the trigger efficiency in data and simulation, to take into account that the derived correction factors differs for the statistical uncertainty and for differences between factors determined from different processes, $W + jets$, $Z + jets$ and $t\bar{t}$ events.

The uncertainties on the lepton and jet energy scales and resolutions are propagated into the uncertainty on E_T^{miss}.

A further contribution to E_T^{miss} also comes from energy deposits that are not associated with any identified physics object; uncertainties on the energy calibration and resolution of the sum of these deposits are also propagated to the uncertainty on E_T^{miss} [135].

5.10.2 Theoretical uncertainties

The limitation of the calculations done in the MC simulation available implies theoretical uncertainties on the analysis procedure; in particular, they affect the normalisation of the diboson and top quark production backgrounds, the shapes of the final discriminant and the signal acceptance. They are more strictly related to the choice of the MC generator, PDFs and parton showers models used, the tuning of the underlying events and any further significant tuning of the MC simulation.

The background contributions coming from diboson and single top-quark processes are estimated from MC simulations and they are normalized to their theoretical cross-sections. For diboson process, the cross-section uncertainty is estimated to be 10% [136]-[115]. In addition, renormalization and factorization scale and PDF variations estimates the uncertainty on the diboson normalization with a further impact of 6%.

An additional contribution from electroweak production, simulated with Madgraph+Pythia8 estimates an increase in the normalization of the diboson background in the VBF analysis by 1.60 (1.85) in the resolved (merged) regime. A conservative uncertainty of 50% is considered on the normalization of the electroweak diboson contribution. The impact on the ggF/DY analysis is negligible. For the cross section of single top-quark process, conservative 20% uncertainty is considered.

Modelling uncertainties on the shapes of the final mass discriminant distributions have been estimated by varying the renormalization/factorization scales, PDF set and α_s values used in the nominal MC samples. The comparison between nominal samples and samples produced with alternative generators estimate the uncertainties due to the choices of generators, parton shower models and event tunes.

The contribution from $V + jets$ and $t\bar{t}$ are constrained using control regions in 1 and 2 lepton channels. For $V + jets$, the nominal Sherpa samples are compared with samples produced using MadGraph5. Moreover, the resummation scale and the CKKW [137]-[138] matching scale in the nominal samples are also varied. The shape systematic uncertainty is typically smaller than 10%, with the Sherpa-Madgraph comparison reaching 25% in the merged ggF WZ untagged signal regions for the $l\nu qq$ channel. For $t\bar{t}$, the default Powheg-Box sample is compared with the alternate $MadGraph5_aMC@NLO$ sample interfaced with Pythia8.230. The shape difference is found to be approximately 4% in the merged signal regions, doubling the value in the resolved analysis. The difference between the Pythia 8.230 using the A14 tune and the alternate Herwig 7.04 [139] using the H7UE set of tuned parameters [140] and the MMHT2014LO PDF set [141] is found between 2 and 5%. The changes from the parameter variations of the nominal generator are less than 5%.

In the 0-lepton channel, no significantly pure region is possible to define to constrint this background. The normalization factors are then extrapolated from 1 lepton ($W + jets$ and $t\bar{t}$) and 2 lepton channel ($Z + jets$), assuming the same normalization factors. Systematic uncertainties on this normalization are obtained by data/prediction double ratio between the default and the alternative MC generator and it is estimated between 10 and 20% for $V + jets$ and up to 30% for $t\bar{t}$. $t\bar{t}$ systematic uncertainty has not been considered on the llqq due to the reduce statistics after selection.

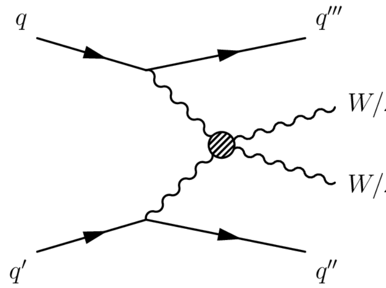
Uncertainties on signal acceptances are estimated because a particular choice of the parton distribution function, QCD scale, the modelling of the initial- and final-state radiations is considered. The PDF uncertainties are estimated by taking the acceptance difference due to internal PDF error sets and the difference between the choice of PDF sets. The effect of the QCD scale uncertainty on the signal acceptance is estimated by varying the factorization and renormalization scales and is applied only to the heavy Higgs model. The uncertainty due to ISR/FSR modelling is studied by varying parameters in the tunes used and applied to the HVT, the RS Graviton and Radion models. These uncertainties, calculated in several resonant mass points, are retrieved for each model, production and decay and considered as conservative. The PDF uncertainties were evaluated to be under 5% for all models. ISR/FSR uncertainties ranges from 2% in the HVT model merged regime to about 11% in the VBFHVT model in resolved regime. The QCD scale for the heavy Higgs model is 1% in ggF production and up to 7% in VBF producion.

5.11 Non-resonant interpretation: experimental signature

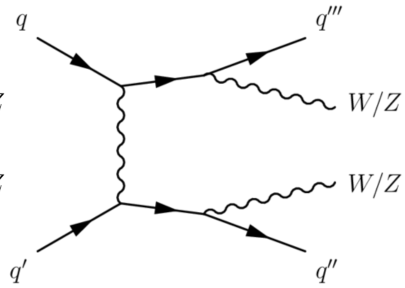
The study of the VBS process provides a relevant test of the electroweak nature of the Symmetry Breaking mechanism 1.4.4 testing the possibility of other sources besides the Higgs mechanism. Theories beyond the SM foresees different coupling in the quartic vertex of the VBS process and deviations from the SM could be observed in this final state.

The VBS process implies an experimental signature with two vector boson and an extra pair of forward jets. The two extra jets represent a crucial piece of the topology to identify the process; the two jets are expected to have high invariant mass and high angular separation, as discussed in Section 1.6.

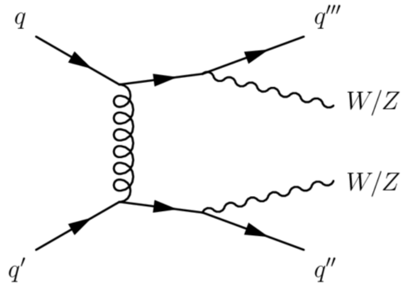
The $VVjj$ process at tree level has both EW (Figure 5.33a 5.33b) and QCD (Figure 5.33c) diagrams The pure EW contributions could be further divided into two components. The first one involves the scattering of two SM bosons, Figure 5.33-a; the scattering occurs trough a quartic gauge coupling (QGC) or a triple gauge vertex involving a Higgs boson or a W/Z boson in the s or t channel; the second one has only EW vertices too, but the two bosons are not produced trough a scattering of two parents bosons.



(a) *EW production via VBS*



(b) *EW production via non-VBS*



(c) *QCD production*

Figure 5.33: Main Feynman diagrams of the EW VVjj production.

The EW non-VBS component cannot be separated from the EW VBS component in a gauge invariant way [142] and contributes significantly to the total cross section. It is therefore included in the signal generation.

5.11.1 Event Selection

The event selection is very similar to the one applied in the resonant case (Section 5.4). The event selection for all channels is summarized in Table 5.12. Further details are given below.

The events collected are categorised firstly according to the leptons multiplicity (0, 1, 2 leptons). Events are categorized into the 0-, 1- and 2-lepton channels depending on the number of selected electrons and muons. In addition to a leptonically decaying candidate V_l, events in all three channels are required to contain a hadronically decaying candidate V_h and two additional small-R jets (referred to as tagging-jets). The V_h candidate could be reconstructed as either two small-R

Selection	0 lep	1 lep	2 lep
Trigger	met triggers	Single-electron triggers Single-muon or met triggers	Single-lepton triggers
Leptons	0 'loose' leptons with $p_T > 7$ GeV	1 'tight' lepton with $p_T > 27$ GeV 0 'loose' leptons with $p_T > 7$ GeV	2 'loose' leptons with $p_T > 20$ GeV GeV 1 lepton with $p_T > 28$ GeV
M^T_{miss}	> 200 GeV	> 80 GeV	–
Multijet removal	$p_T^{miss} > 50$ GeV $\Delta\Phi(\vec{E_T^{miss}}, \vec{p_T^{miss}}) < \pi/2$ $min[\Delta\Phi(\vec{E_T^{miss}}, small - Rjets] > \pi/6$ $\Delta\Phi(\vec{E_T^{miss}}, V_{had}) > \pi/9$	–	–
$m_{\ell\ell}$	–	–	$83 < m_{ee} < 99$ GeV $(-0.0117 \times p_T^{\mu\mu} + 85.63 \text{ GeV}) < m_{\mu\mu} < (0.0185 \times p_T^{\mu\mu} + 94 \text{ GeV})$
Small-R jets	$p_T > 20$ GeV if $\lvert\eta\rvert < 2.5$, and $p_T > 30$ GeV if $2.5 < \lvert\eta\rvert < 4.5$		
Large-R jets	$p_T > 200$ GeV, $\lvert\eta\rvert < 2$		
$V_{had} \to J$ $V_{had} \to jj$	V boson tagging, $min(\lvert m_J - m_Z\rvert, \lvert m_J - m_W\rvert)$ $64 < m_{jj} < 106$ GeV, jj pair with $min(\lvert m_{jj} - m_Z\rvert, \lvert m_{jj} - m_W\rvert)$, leading jet with $p_T > 40$ GeV		
Tagging-jets	$j \notin Vhad$, not b-tagged, $\Delta R(J, j) > 1.4$ $\eta_{TagJet1} \cdot \eta_{TagJet2} < 0$, $m_{jj}^{Tag} > 400$ GeV, $p_T > 30$ GeV		
Num. of b-jets	–	0	–

Table 5.12: Summary of the event selection in the 0-, 1- and 2-lepton channels.

jets ($V \to jj$) in a resolved selection, or one large-R jet ($V \to J$) in a merged selection. The event selection criteria has been optimised to maximize the sensitivity of the analysis defining nine non-overlapping distinct signal regions (SR): one for each of the three lepton channels and three types of V_h selections (resolved, and low- and high-purity merged).

Signal events in the 0 lepton channel are characterised by a hadronically decaying V boson recoiling against a large amount of missing transverse momentum due to either a $Z \to \nu\nu$ decay or a $W \to l\nu$ decay, when the lepton is outside the acceptance of the detector. An initial selection is made by requiring $E_T^{miss} > 200\ GeV$, and rejecting events with electrons or muons passing the loose quality requirements. The multijet background is suppressed using a requirement on the value of the track-based missing transverse momentum, $p_T^{miss} > 50\ GeV$. Three further angular selection criteria are reported in table 5.12.

The 1 lepton channel is typical of a leptonically decaying W boson. The candidates are selected by requiring one isolated lepton satisfying the "tight" criteria with $p_T > 27\ GeV$. Events are required to have $E_T^{miss} > 80\ GeV$, and must not have any additional "loose" leptons. In order to reconstruct the invariant mass of the WV system, needed later to construct the multivariate discriminant, the neutrino momentum four-vector is reconstructed by imposing a W boson mass constraint on the lepton-neutrino system. The neutrino transverse momentum components are set equal to the missing transverse momentum of the event and the unknown z-component of the momentum (p_z) is obtained from the resulting quadratic equation. The p_z is chosen as either the smaller, in absolute value, of the two real solutions or, if the solution is complex, its real part.

In the 2 leptons channel, the $Z \to ll$ candidates are identified by requiring two isolated same-flavor leptons satisfying the "loose" criteria. The leading (subleading) lepton must satisfy $p_T > 28\ (20)\ GeV$. The dilepton invariant mass is required to be consistent with that of the Z boson: $83 < m_{ee} < 99\ GeV$ in the case of electrons and $(-0.0117 \cdot p_T^{\mu\mu} + 85.63\ GeV) < m_{\mu\mu} < (0.0185 \cdot p_T^{\mu\mu} + 94\ GeV)$ in the case of muons. The p_T-dependent requirement on $m_{\mu\mu}$ recovers the selection efficiency at high $p_T^{\mu\mu}$, which would otherwise fall due to the degraded dimuon invariant mass resolution [53], as already discussed.

The merged selection is applied as the first step in identifying a V_h candidate. If an event is not selected, then the resolved selection is used. Merged selection events are required to have at least one large-R jet. Next the boson tagging discussed in Section 3.3.3 is applied to identify the $V \to qq$ decays. Two SRs are defined, one for events passing the 50% working point of the boson tagging requirement (HP) and the other for events failing the 50%, but passing the 80% working point requirement (LP). The large-R jets are required to satisfy either W or Z boson tagging. If multiple V_h candidates are selected, the one minimizing the $min(|m_J - m_Z|, |m_J - m_W|)$ is selected.

The resolved selection events are required to have two small-R signal jets with a dijet invariant mass inside the $m_{W/Z}$ window: $64 < m_{jj} < 106\ GeV$. If multiple V_h candidates are selected, the one minimizing $min(|m_{jj} - m_Z|, |m_{jj} - m_W|)$ is used. At least one of the jets forming the selected V_h candidate must have $p_T > 40\ GeV$, in order to improve the separation between the signal and the background; otherwise the event is not selected.

After selecting the V_h candidate, tagging-jets are selected from the remaining small-R jets that fail the b-tagging described in Section 3.3.4 for the resolved regime and from all small-R jets with $\Delta R(J, j) > 1.5$ for the merged one. Tagging-jets are required to be in opposite hemispheres, $\eta_{TagJet1} \cdot \eta_{TagJet2} < 0$, and the invariant mass of the two tagging-jets must satisfy $m_{jj}^{Tag} > 400\ GeV$ 5.11.3. Both tagging-jets must have $p_T > 30\ GeV$ in order to suppress the contribution from pileup interactions, otherwise the event is rejected.

5.11.2 Control regions

The $W+jets$ and $t\bar{t}$ production are the dominant backgrounds for the 1 lepton channel, the $Z+jets$ is the dominant one for the 2 lepton channel; while they contribute both in the 0 lepton channel. Smaller background contributions for the 1 lepton channel arise from multi-jets background. Single-top and QCD-induced diboson production is a small background for all three lepton channels. The background contributions are estimated using a combination of MC and data-driven techniques. The shapes of kinematic variable distributions are taken from MC simulations in all cases except for the multi-jet background in the 1 lepton channel.

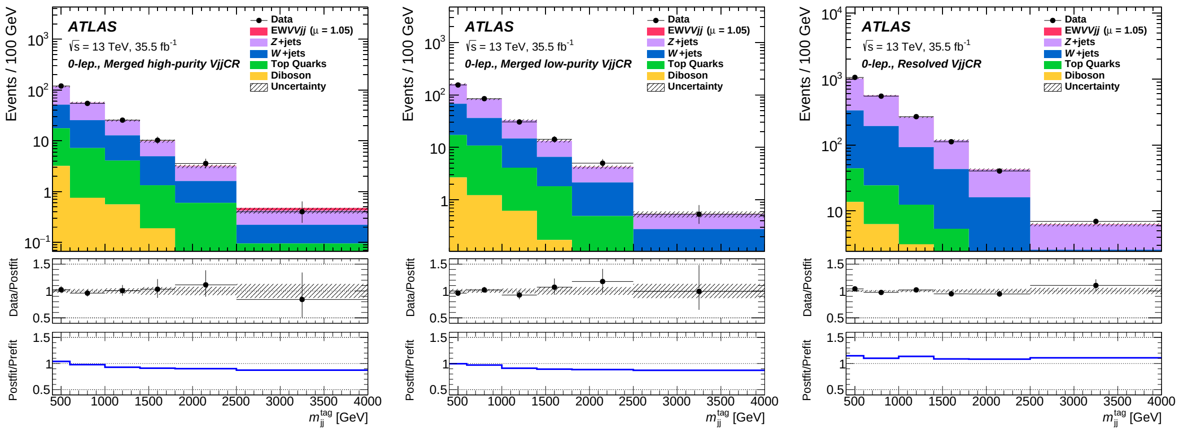
A $Z + jets$ control region (ZCR) is defined for each of the three SRs in the 2 leptons channel by reversing the m_J or m_{jj} requirement. For the merged selection, the leading large-R jet mass is required to be outside the large-R jet mass window of the 80% working point of the W/Z boson tagging. These CRs are dominated by the $Z + jets$ contribution, with a purity higher than 95% in all regions.

Three $W + jets$ control regions (WCRs) are defined from events satisfying the 1 lepton signal region selection except for the invariant mass requirement of the V_h candidate, similar to the ZCRs. Approximately 86% and 77% of the selected events are from $W + jets$ production in the merged and resolved categories of the 1 lepton channel, respectively. The remaining events are primarily from $t\bar{t}$ production.

The three $t\bar{t}$ control regions (TopCRs) consist of events satisfying the signal region selection of the 1 lepton channel except for the $b - jet$ requirement, which is inverted. These CRs are dominated by $t\bar{t}$ production, with a purity of 79% and 59% for merged and resolved categories respectively, and the remainder are from single-top, $V + jets$ or diboson production, for both the merged and the resolved event topologies.

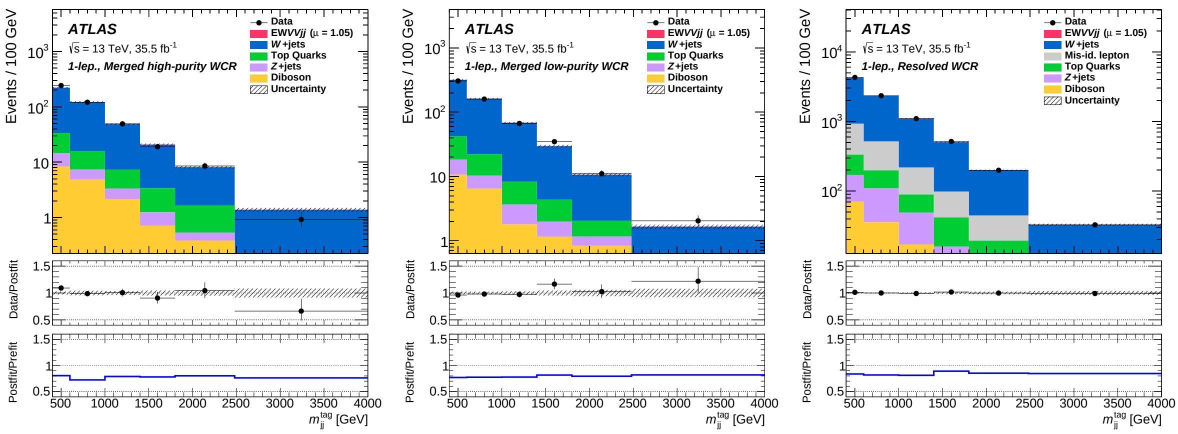
In the 0 leptons channel, it is not possible to define pure control regions for $W + jets$, $Z + jets$ and $t\bar{t}$ processes, thus events falling into the mass sideband regions of the V_h, similar to WCRs and ZCRs, form three different CRs (referred to as VjjCR), one for each of the corresponding SRs.

The m_{jj}^{tag} spectra of simulated $W + jets$ ($Z + jets$) events are not well modelled by the MC simulation in the WCRs (ZCRs) for the three V_h selections in the 1-lepton (2-lepton) channel. A data-driven procedure is applied to the simulated $W + jets$ and $Z + jets$ events to correct for this shape mismodeling. Reweighting factors are derived from WCRs and ZCRs as a function of m_{jj}^{tag}, and applied to all SRs and CRs (for 0-, 1-, and 2-lepton regions) in the MC simulation of $W + jets$ and $Z + jets$ events, respectively. The non-$W + jets$ ($Z + jets$) contributions are subtracted from the spectra in data. Then the re-weighting factors as a function of m_{jj}^{Tag} are determined by performing a linear fit to the ratios of data to simulation in the control regions. The re-weighting is done separately for the merged and resolved analyses. For $W + jets$, the re-weighting factor ranges from 1.016 (1.024) at $m_{jj}^{tag} = 400\ GeV$ to 0.47 (0.53) at $m_{jj}^{tag} = 3000\ GeV$ in the resolved (merged) analysis. For $Z + jets$, the re-weighting factor ranges from 1.071 (1.062) at $m_{jj}^{tag} = 400\ GeV$ to 0.42 (0.36) at $m_{jj}^{tag} = 3000\ GeV$ in the resolved (merged) analysis.



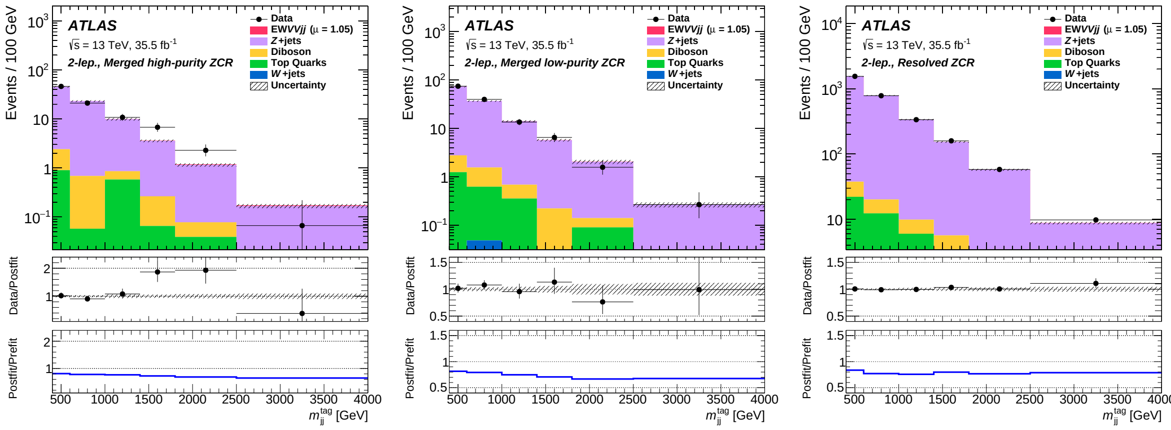
(a) (b) (c)

Figure 5.34: m_{jj}^{Tag} distribution in the CR for the 0 leptons channel.



(a) (b) (c)

Figure 5.35: m_{jj}^{Tag} distribution in the CR for the 1 leptons channel.



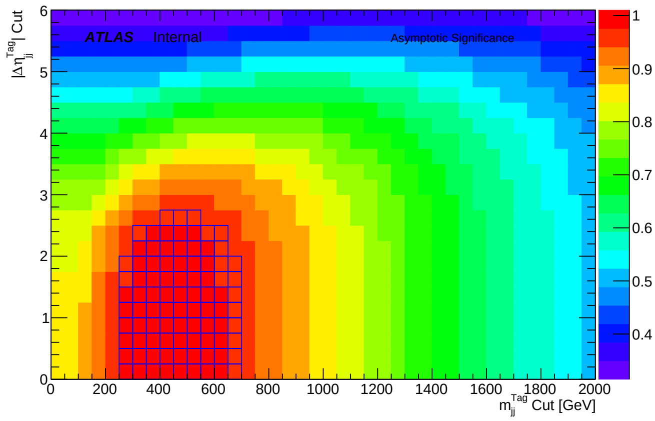
(a) (b) (c)

Figure 5.36: m_{jj}^{Tag} distribution in the CR for the 2 leptons channel.

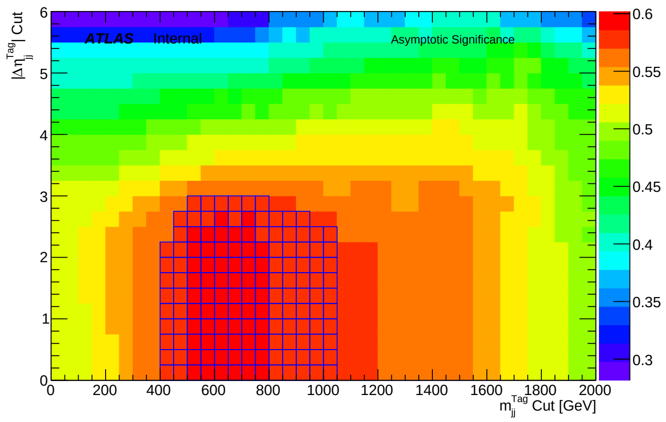
5.11.3 Optimisation of VBS tagging jets selection

A dedicated study has been perfomed in order to optmize the cut selection of the VBS tag jets.

The expected asymptotic significance has been evaluated varying the cuts on the two more discriminating variables, m_{jj}^{Tag} and $|\Delta\eta_{jj}^{Tag}|$. Two dimensional maps, reported in Figure 5.37, show the estimator $\epsilon_{VBS} \cdot (1 - \epsilon_{bkg})$ for the resolved and merged regime. An optimal area around the maximum of the map in which the significance variation is less then the 5% has been considered. The estimator does not improve significantly if a $|\Delta\eta_{jj}^{Tag}|$ cut is applied. Thus, the optimal cut is applied as $m_{jj}^{Tag} > 400\ GeV$.



(a)

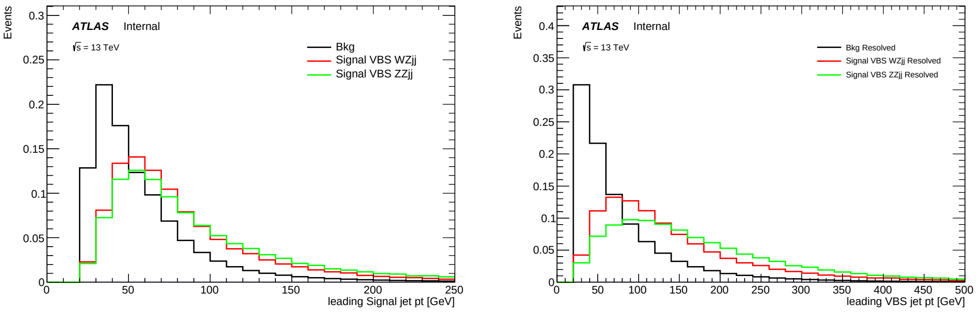


(b)

Figure 5.37: 2-dimensional maps of the estimator $\epsilon_{VBS} \cdot (1 - \epsilon_{bkg})$ varying the cuts on the invariant mass and the η separation of the tag jets.

5.11.4 Jet p_T cut optimisation

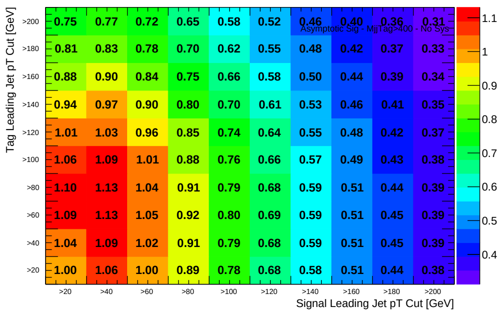
A study about the p_T cut of the leading and sub-leading jets has been performed in order to maximize the sensitivity of the analysis. Default cuts of 20 GeV were applied. In Figure 5.38 the p_T distributions for the Signal Leading Jet and for the Tag Leading Jet (resolved regime) are showed.



(a) *Signal Leading Jet pT* (b) *Tag Leading Jet pT*

Figure 5.38: p_T distributions (resolved regime).

Two dimensional maps for the asymptotic significance improvement has been obtained varying the cuts over the leading signal and tag jets p_T. The optimal cuts are around 40 GeV for the signal leading jet and of 60 GeV for the tag leading with an expected improvement of 10% (Figure 5.39).

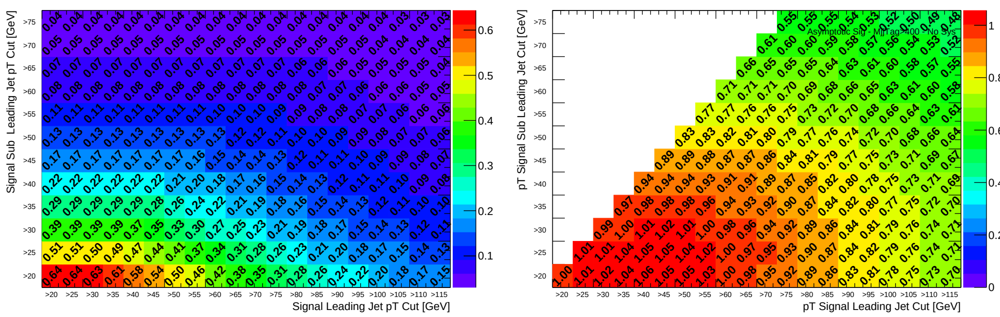


(a)

Figure 5.39: Map varying the p_T cut values for the signal leading and tag leading jets and considering the case $M_{jj}^{Tag} > 400$ GeV.

A study has been performed also attempting to optimize the sub-leading jet p_T. Maps of the asymptotic significance gain varying the cut values for the Signal leading jet p_T and for the signal sub-leading jet p_T were produced considering different cut values on the M_{jj}^{Tag} variable; as an example, the case related to $M_{jj}^{Tag} > 400$ GeV is shown in Figure 5.40.

Cuts are applied on the p_T of the Tag Jets pair and of the Signal Jets pair; the values chosen were of 40 GeV for the Leading Signal Jet and of 30 GeV on both Leading and Sub-Leading Tag Jets, that optimize the expected significance of the signal.



(a) (b)

Figure 5.40: Efficiency map (a) and expected asymptotic significance improvement maps in the nominal case (b) for the signal jets.

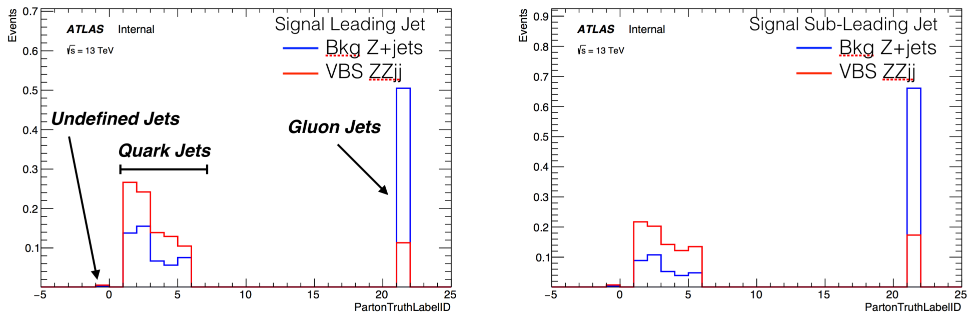
$M_{jj}^{Tag} cut$	SignalJet1 pT cut	SignalJet2 pT cut	Signal Efficiency	Significance	Significance (5% Sys)	Signficance (10% Sys)
0	40	20	0.90	0.90	0.17	0.10
0	50	25	0.62	0.90	0.21	0.13
200	20	20	0.87	0.96	0.19	0.11
200	40	20	0.79	1.03	0.23	0.13
200	50	25	0.55	1.02	0.29	0.17
400	20	20	0.64	1.01	0.26	0.15
400	40	20	0.58	1.07	0.30	0.18

Table 5.13: Signal efficiency and expected asymptotic significance using different $M_{jj}^{Tag} cut$ and p_T cut values.

5.11.5 Tracks multiplicity for QG identification

The VBS final state is expected to be with at least four jets quark-initiated. The main expected backgrounds *Z+jets*, instead, could have a significant fraction of events with gluon jets in the final state. This concept motivate the use of quark to gluon identification in the non-resonant interpretation of diboson analyses.

In Figure 5.41 the *Parton Truth Label ID* distributions of the two Signal Jets is showed. As expected, the background sample is characterized by a greater gluon component respect to the VBS signal sample, in the table 5.14 all the fraction of truth parton initiated jets are reported.



(a) (b)

Figure 5.41: Parton Truth Label ID distribution of the Signal Leading Jet (a) and of the Signal Sub-Leading Jet (b).

	Signal VBS ZZjj			Bkg Z+jets		
	Undefined	Quark	Gluon	Undefined	Quark	Gluon
Signal Leading Jet	-	88%	12%	-	49%	51%
Signal Sub-Leading Jet	1%	82%	17%	-	34%	66%

Table 5.14: Quark and Gluon compositions of the Signal Jets in the resolved regime.

As discussed in Section 3.3.5, at the present state a simple 1-dimensional quark/gluon tagger has been implemented using the tracks multiplicity of the jets. The current version has been calibrated in a specific kinematic range of the jets:

$$p_T > 50\ GeV \qquad |\eta| < 2.1 \tag{5.7}$$

Given the kinematic distributions of the selected jets, only a relative fraction of the them fall inside the acceptance of the tagger. Figure 5.42 shows the p_T and η distributions of the two Signal Jets; the green areas represent the fraction of jets that belong to the acceptance range of the tagger. The fraction of events in which at least one or both two Signal Jets are in the acceptance range of the tagger is shown in table 5.15.

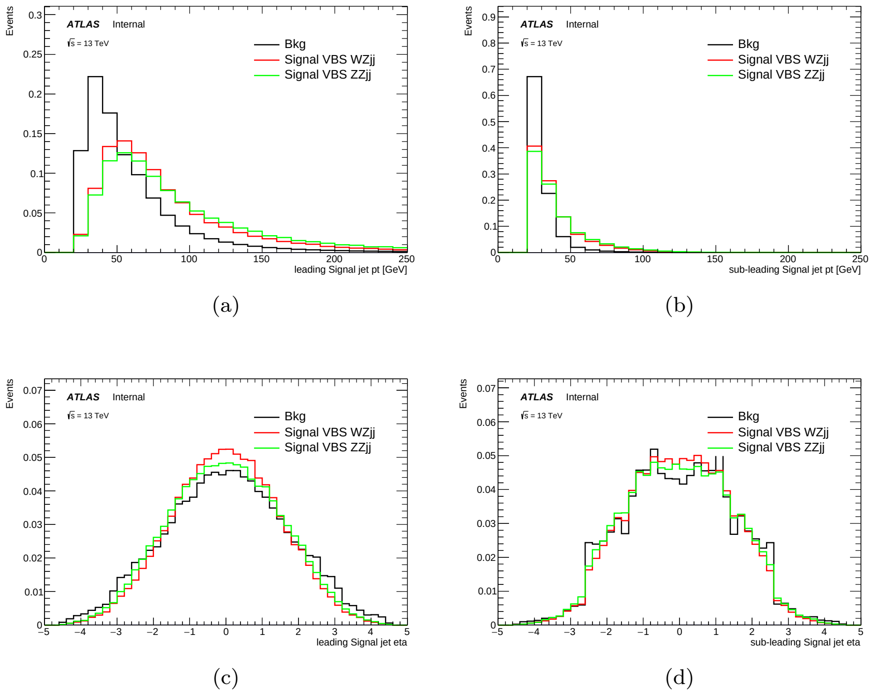
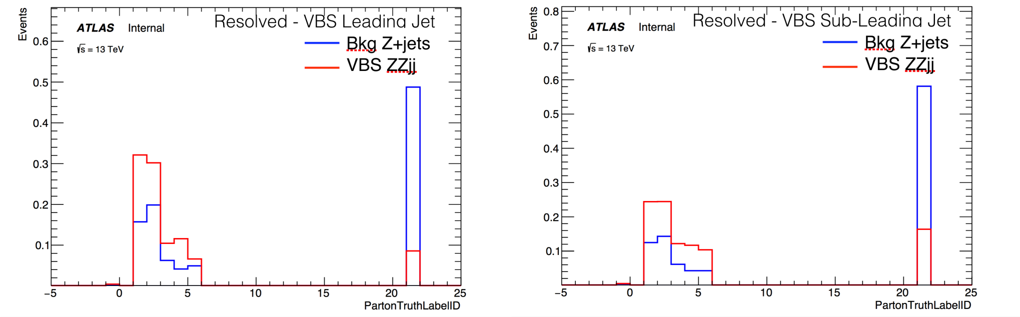


Figure 5.42: pT and η distributions of the Signal Leading Jet (a)-(c) and of the Signal Sub-Leading Jet (b)-(d).

	Signal VBS ZZjj	Signal VBS WZjj	Bkg
1 jet inside QG acceptance range	68%	67%	40%
2 jets inside QG acceptance range	18%	15%	3%

Table 5.15: Fraction of events in which at least one or both two Signal Jets are in the acceptance range of the tagger.

The same study was performed for the Tag Jets pair in both resolved and merged regime.
First, in Figure 5.43 the truth composition of the Tag Leading and of the Tag Sub-Leading jets is shown. As for the signal jets, the gluon component is larger for the background sample (50%). It is order $\sim 10-20\%$ for signal sample (Table 5.16).



(a) (b)

Figure 5.43: Parton Truth Label ID distribution of the Tag Leading Jet (a) and of the Tag Sub-Leading Jet (b) in the resolved regime.

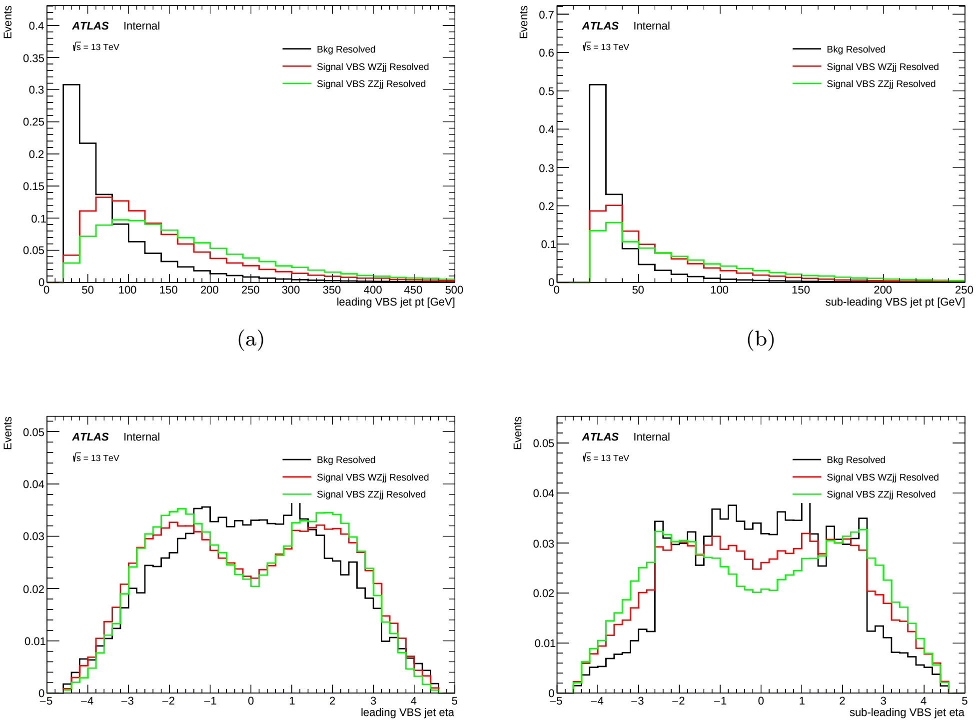
	Signal VBS ZZjj			Bkg Z+jets		
	Undefined	Quark	Gluon	Undefined	Quark	Gluon
Tag Leading Jet	-	91%	9%	-	51%	49%
Tag Sub-Leading Jet	-	83%	17%	-	42%	58%

Table 5.16: Quark and Gluon compositions of the Tag Jets in the resolved regime.

p_T and η distributions for the Tag Leading Jet ((a)-(c)) and for the Tag Sub-Leading Jet ((b)-(d)) are reported in Figure 5.44; According to the tagger acceptance, only a fraction of jets fall inside the acceptance allowed for the quark or gluon jets identification; the fraction of events in which at least one jet or both two tag jets are taggable are evalueted in table 5.17.

	Signal VBS ZZjj	Signal VBS WZjj	Bkg
1 jet inside QG acceptance range	70%	66%	44%
2 jets inside QG acceptance range	21%	18%	9%

Table 5.17: Fraction of events in which at least one or both two Tag Jets in the resolved regime are in the acceptance range of the tagger.



(a) (b)

(c) (d)

Figure 5.44: pT and η distributions of the Tag Leading Jet (a)-(c) and of the Tag Sub-Leading Jet (b)-(d) in the resolved regime.

	Signal VBS ZZjj			Bkg Z+jets		
	Undefined	Quark	Gluon	Undefined	Quark	Gluon
Tag Leading Jet	1%	86%	13%	-	51%	49%
Tag Sub-Leading Jet	1%	75%	24%	-	37%	63%

Table 5.18: Quark and Gluon compositions of the Tag Jets in the merged regime.

	Signal VBS ZZjj	Signal VBS WZjj	Bkg
1 jet inside QG acceptance range	73%	71%	75%
2 jets inside QG acceptance range	22%	20%	20%

Table 5.19: Fraction of events in which at least one or both two Tag Jets in the merged regime are in the acceptance range of the tagger.

Upper limit on the expected significance gain coming from a q/g tagging can be obtained using truth information and removing events with gluon jets. This study indicates a maximum gain of 20% in the merged regime while of 40% in the resolved regime, as reported in table 5.20.

	σ'/σ
Merged	1.19
Resolved	1.44

Table 5.20: Significance gain in the merged and in the resolved regimes using the Parton Truth Label ID.

This kind of selection represents an ideal tagging of the jets and therefore the gain we obtained represents the best improvment that could be obtained. Significance gain can be, in principle, enhanced by a larger eta-pT acceptance of the qg tagger. The impact of the extending of the acceptance range is showed in Figure 5.45.

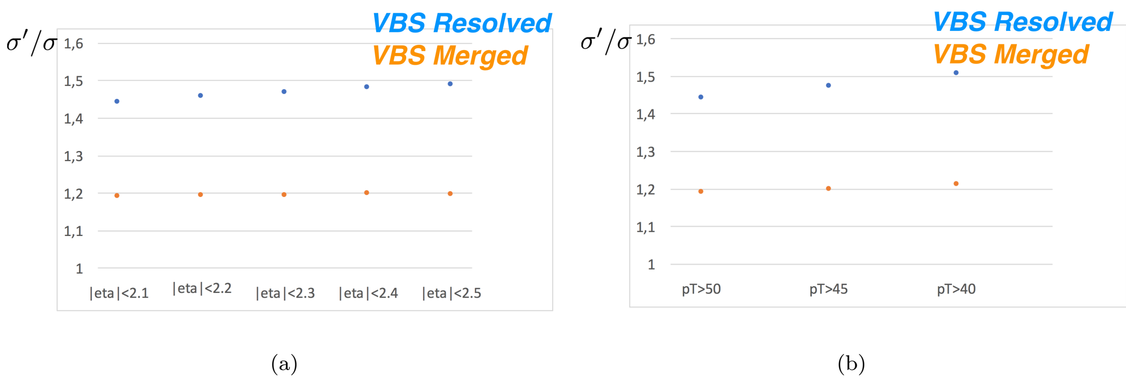


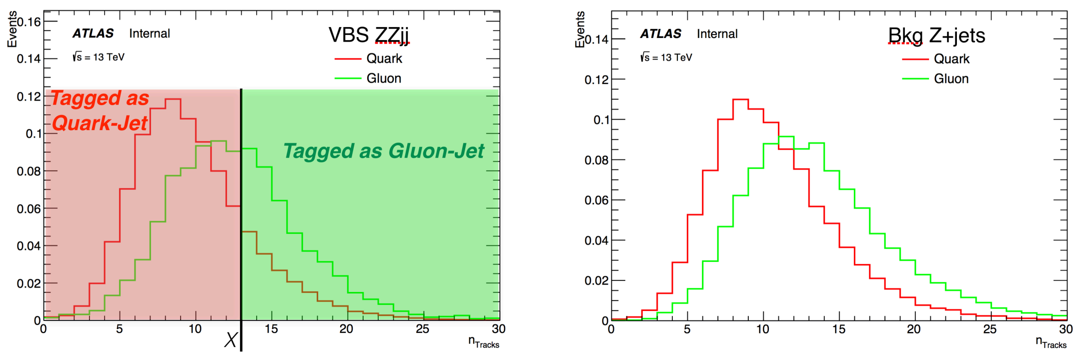
Figure 5.45: Significance gain as a function of q/g acceptance on $|\eta|$ (a) and p_T (b).

As expected, the gluon-jets are characterized by a multiplycity of charged particles greater then quark-jets.

The tagger is based on a simple cut applied on the tracks variable: all jets with a number of tracks less then X are tagged as *quark-jets* while all jets with a number of tracks greater then X are tagged as *gluon-jets*. This is schematically represented in Figure 5.46-a.

In order to show the impact of a quark vs gluon jet identification based on the number of tracks variable, efficiency curve varying the cut value of the number of tracks variable were obtained. In Figure 5.47-(a)-(b) the blue curve is referred to the efficiency selection of truth quark jets tagged as quark jets while the red curve is referred to the mistagging of gluon jets, that is the truth gluon jets tagged as quark jets.

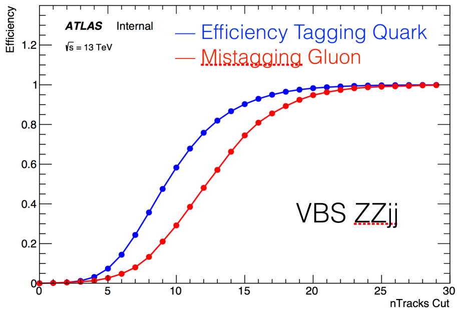
Then, the efficiency curves of Figure 5.47-(a)-(b) have been combined in the ROC curve of Figure 5.47-(c).



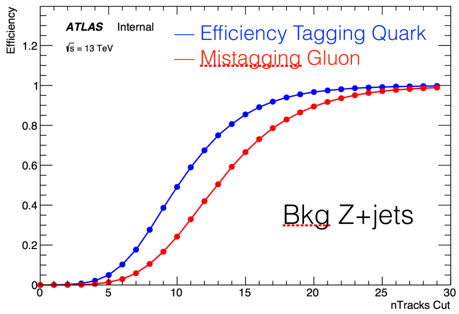
(a) (b)

Figure 5.46: Distribution of the number of tracks of the Leading Signal Jet for the Signal VBS ZZjj (a) and the background sample (b); the red histrogram represents the truth quark component while the green one the gluon component.

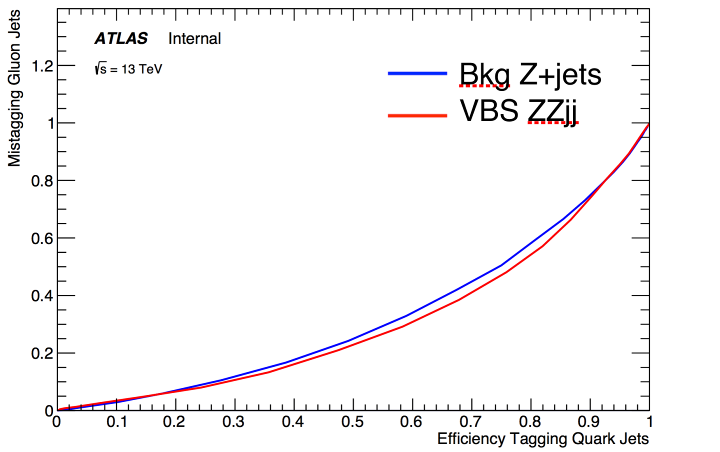
A scan on the value of the cut for the number of tracks variable has been performed, the cut has been applied to the signal and tag jets in resolved regime while only to tag jets in the merged regime. The results is showed in Figure 5.48. A non negligible significance gain is obtained both in the resolved and merged regime. The optimal cut is around 14 and the corresponding gain are of $\sim 10\%$ in the resolved regime and of $\sim 5\%$ in the merged regime.



(a)

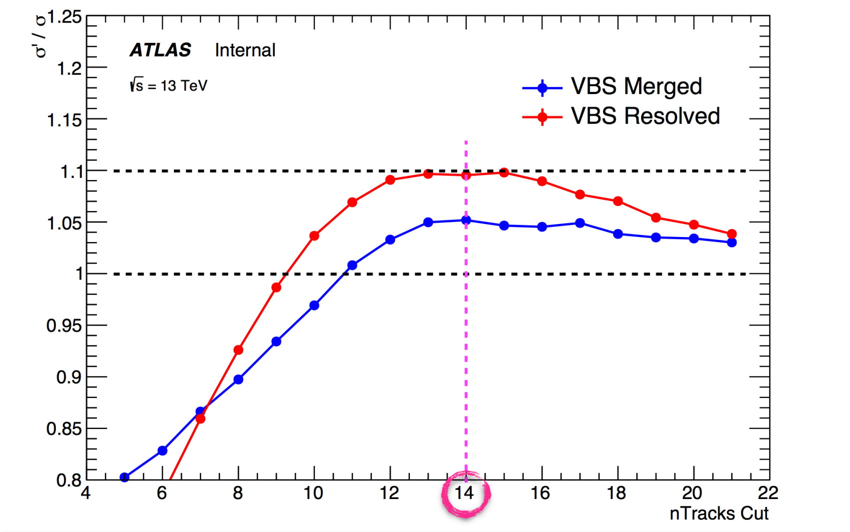


(b)



(c)

Figure 5.47: Efficiency curves (a)-(b) and the ROC curve (c) obtained varying the cut on the $nTracks$ variable for the Leading Signal Jet.



(a)

Figure 5.48: Improvement on the expected significance as a function of the cut on the n_{Tracks}.

Variable	0-lepton	1-lepton	2-lepton
m_{jj}^{Tag}	✓	–	✓
$\Delta\eta_{jj}^{Tag}$	–	–	✓
$p_T^{TagJet2}$	✓	✓	✓
m_J	✓	–	–
$D_2^{(\beta=1)}$	✓	–	✓
E_T^{miss}	✓	–	–
$\Delta\Phi(E_T^{\vec{miss}}, J)$	✓	–	–
η_l	–	✓	–
$n_{TrkJets}$	✓	–	–
ζ_V	–	✓	✓
m_{VV}	–	–	✓
p_T^{VV}	–	–	✓
m_{VVjj}	–	✓	–
p_T^{VVjj}	–	–	✓
$TracksWidth_{TagJet1}$	✓	–	–
$TracksWidth_{TagJet2}$	✓	–	–

Table 5.21: Variables used for the BDT discriminant in the merged analysis category of each lepton channel.

5.12 Machine Learning approach

A multivariate approach has been used after the events selection in order to maximise the separation between the signal and the background; indeed, a final discriminant is not naturally provided from the kinematics, since the invariant mass of the diboson system is not expected to peak, as for the resonant case, and the invariant mass of the extra di-jets system is not well modelled as described in the previous section.

The analysis uses the Toolkit for Multivariate Analysis (TMVA) provided in the ROOT software framework [89]. A Boost Decision Tree (BDT) algorithm has been developed. A single BDT has been defined and trained in each lepton channel and in each regime of the hadronic boson reconstruction (Merged HP and Merged LP have been merged together in order to enhance the statistics) in order to take into account the different phase space selected; so, six different BDT have been considered: (0, 1, 2) leptons x (resolved, merged) regimes.

The most discriminating variables for the VBS signal are used; several variables related to the Tag Jets pair and the Signal Jets pair are used and of the full diboson system; furthermore, q/g informations, such as the tracks multiplicity and the calorimetric width are used since the discrimination power exploited in the previous Section 5.11.5. A summary of the variables used as BDT inputs are reported in tables 5.21-5.22.

Variable	0-lepton	1-lepton	2-lepton
m_{jj}^{Tag}	✓	–	✓
$\Delta\eta_{jj}^{Tag}$	–	–	✓
$p_T^{TagJet1}$	✓	✓	–
$p_T^{TagJet2}$	✓	✓	✓
$\Delta\eta_{jj}$	✓	✓	✓
p_T^{j1}	✓	–	–
p_T^{j2}	✓	✓	✓
$TracksWidth_{j1}$	✓	✓	✓
$TracksWidth_{j2}$	✓	✓	✓
n_{Tracks}^{j1}	–	✓	✓
n_{Tracks}^{j2}	–	✓	✓
$TracksWidth_{TagJet1}$	✓	✓	✓
$TracksWidth_{TagJet2}$	✓	✓	✓
$n_{Tracks}^{TagJet1}$	–	✓	✓
$n_{Tracks}^{TagJet2}$	–	✓	✓
$n_{TrackJets}$	✓	–	✓
$n_{TrackJets}^{extra}$	✓	–	–
E_T^{miss}	✓	–	–
η_l	–	✓	–
$\Delta R(l,\nu)$	–	✓	–
ζ_V	–	✓	✓
m_{VV}	–	–	✓
m_{VVjj}	–	✓	–

Table 5.22: Variables used for the BDT discriminant in the resolved analysis category of each lepton channel analysis.

5.13 Fiducial cross section definition

The cross section of the EW signal has been measured in a region of the kinematic phase space close to the acceptance of the detector in each of the leptons channel. Only stable final-state particles are considered in the definition of the phase space. [143] Leptons produced in the decay of a hadron or its products are not considered in the charged lepton requirement of the fiducial phase space. The selection of the fiducial phase space is summarised in the table 5.23

<table>
<tr><th colspan="3">Object selection</th></tr>
<tr><td>Leptons</td><td colspan="2">$p_T > 7$ GeV, $|\eta| < 2.5$</td></tr>
<tr><td>Small-R jets</td><td colspan="2">$p_T > 20$ GeV if $|\eta| < 2.5$, and $p_T > 30$ GeV if $2.5 < |\eta| < 4.5$</td></tr>
<tr><td>Large-R jets</td><td colspan="2">$p_T > 200$ GeV, $|\eta| < 2.0$</td></tr>
<tr><th colspan="3">Event selection</th></tr>
<tr><td rowspan="3">Leptonic V selection</td><td>0-leptons</td><td>Zero leptons, $p_T^{\nu\nu} > 200$ GeV</td></tr>
<tr><td>1-leptons</td><td>One lepton with $p_T > 27$ GeV, $p_T^{\nu} > 80$ GeV</td></tr>
<tr><td>2-leptons</td><td>Two leptons, with leading (subleading) lepton $p_T > 28$ (20) GeV
$83 < m_{\ell\ell} < 99$ GeV</td></tr>
<tr><td rowspan="2">Hadronic V selection</td><td>Merged</td><td>One large-R jet, $min(|m_J - m_Z|, |m_J - m_W|)$
$64 < m_J < 106$ GeV</td></tr>
<tr><td>Resolved</td><td>Two small-R jets, $min(|m_{jj} - m_Z|, |m_{jj} - m_W|)$
$p_{\mathrm{T}}^{j_1}$ >40 GeV, $p_{\mathrm{T}}^{j_2}$ >20 GeV
$64 < m_{jj} < 106$ GeV</td></tr>
<tr><td>Tagging-jets</td><td colspan="2">Two small-R non-b jets, $\eta_{TagJet1} \cdot \eta_{TagJet2} < 0$, highest m_{jj}^{tag}
$m_{jj}^{tag} > 400$ GeV, $p_{\mathrm{T}}^{\mathrm{tag},j_{1,2}} > 30$ GeV</td></tr>
<tr><td rowspan="3">Number of b-jets</td><td>0-leptons</td><td>–</td></tr>
<tr><td>1-leptons</td><td>0</td></tr>
<tr><td>2-leptons</td><td>–</td></tr>
</table>

Table 5.23: Fiducial phase-space definitions used for the measurement of electroweak VVjj production.

Chapter 6

Analysis Results

In this chapter the final results of the analyses are presented. The unblinded data used to search for new diboson resonances are showed. The full Run-II dataset corresponding to an integrated luminosity of 139 fb^{-1} has been used. No evidence of new signals are present so exclusion limit on the production of new heavy resonances are presented. Moreover, the results for the non-resonant interpretation are showed and the cross section measurement is reported.

6.1 Results of the search for new diboson resonances

In this section a summary of the results obtained studying the diboson final states are presented. The statistical procedure adopted to constraint the SM background predictions using the CRs is described. The data agreement with the expected background is showed and, finally, the limits on the productions of new resonances in the benchmark models obtained from the observed data are derived.

6.1.1 Statistical Procedure

In the search for new diboson resonances a complex statistical analysis have been performed to interpret the results.

A profile-likelihood-ratio test statistic is used to measure the compatibility of the background-only hypothesis with the observed data and to test the signal-plus-background hypothesis for the production of a heavy resonance X, with its production cross section, $\sigma(pp \to X)$ as the parameter of interest. The statistical analysis framework is detailed described in [144, 145, 146].

The different lepton channels have different discriminants: the transverse mass m_T is used in the 0-lepton channel, the $m_{l\nu J}$ or $m_{l\nu jj}$ in the 1-lepton and m_{llJ} or m_{lljj} in the 2-lepton one. Maximum likelihood fits are performed considering the binned distributions of the final discriminants in data. The analysis discriminating distributions are arranged as the product of channel, category, regime and region. Here channel refers to the analysis with different number of leptons in the final states (0-lepton, 1-lepton and 2-lepton); category refers to the selection for VBF or ggF signal; regime refers to the merged and resolved jet analysis objects used in selection; and region refers to signal (SR) and control regions (CR) which normally consist of subregions of different properties. A summary of the number of regions considered in the fit is reported in table 6.1.

The normalizations of the $V + jets$ and $t\bar{t}$ contributions are left float in these fits. Systematic uncertainties, described in 5.10, and their correlations are incorporated into the fits with nuisance parameters, where each has been given a prior distribution based on individual studies or it is allowed to float freely, constrained simultaneously by the SRs and CRs.

Two different fits are performed; one on to the $WW + ZZ$ searches and the second one to the WZ search. The $WW + ZZ$ fit includes all the SRs foreseen for the WW and ZZ final states (compatible with spin-0 and spin-2 resonances) and the related CRs, while the WZ includes all the SRs built for the WZ final state (spin-1 resonances) and the related CRs. Each fit contains regions coming from each individual lepton channel. Separate fits are performed for ggF/DY and VBF production modes and for different resonance mass hypotheses, but including SRs in both ggF/DY and VBF categories.

VV	Signal regions						Control regions		
	VBF category			ggF/DY category			$W+jets$	$Z+jets$	$t\bar{t}$
final state	0-lep.	1-lep.	2-lep.	0-lep.	1-lep.	2-lep.			
WW	0	3	0	0	3	0	6	–	6
ZZ	2	0	3	4	0	6	–	9	–
WZ	2	3	3	2	6	3	9	6	9

Table 6.1: Numbers of signal regions in each leptonic channel and of control regions for different diboson final states.

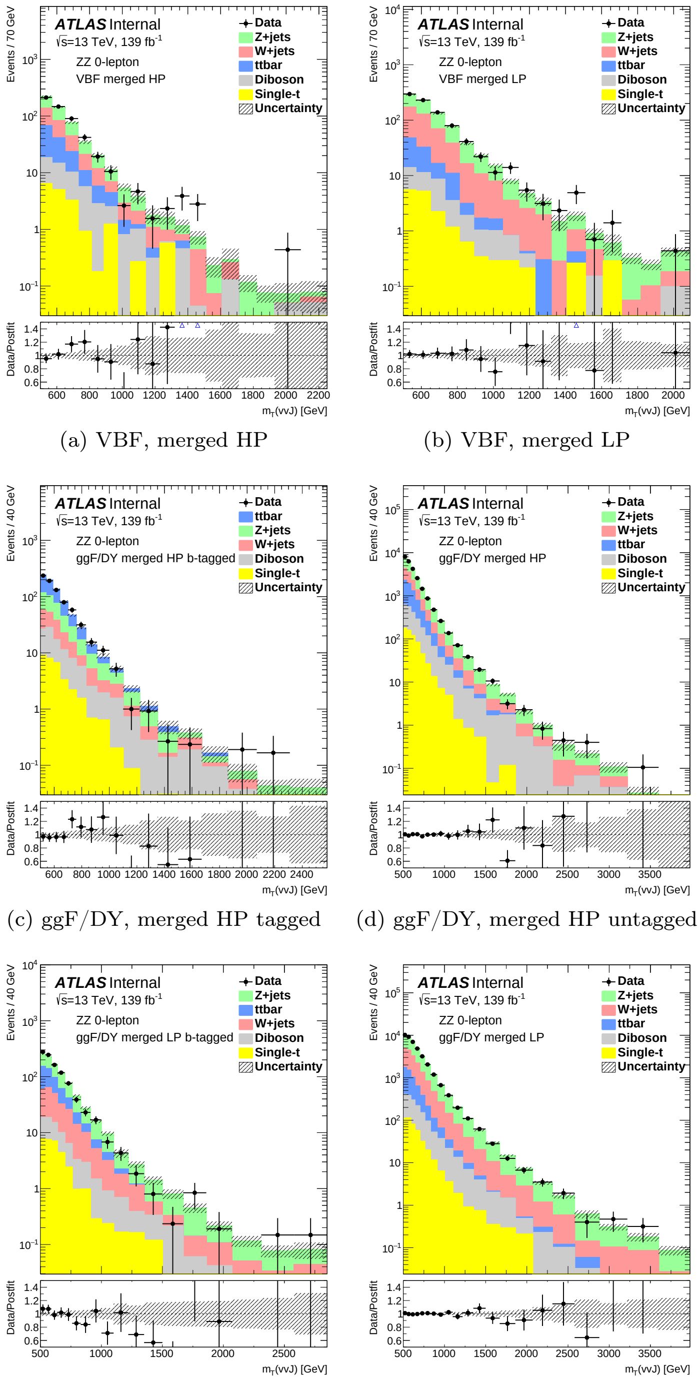
6.1.2 Data and background comparison

The fit is first done in the background-only hypothesis to test the compatibility of the data with the background expectations. The observed distribution of the final discriminants are found to be reasonably well described by their estimated background contributions.

As examples, the data are compared with the expected backgrounds from the WW + ZZ fit in Figure 6.1 for the m_T distributions of the 0-lepton channel, in Figure 6.2 for the $m_{l\nu J}$ / $m_{l\nu jj}$ distributions of the 1-lepton channel, and in Figure 6.3 for the m_{llJ} / m_{lljj} distributions of the 2-lepton channel. Table 6.2 shows the post-fit estimated background event yields from different sources in all WW and ZZ SRs compared with the numbers of the observed events in data. A similar level of agreement is obtained for the WZ fit. No significant excess over the estimated background is observed in the mass distributions of all SRs in data.

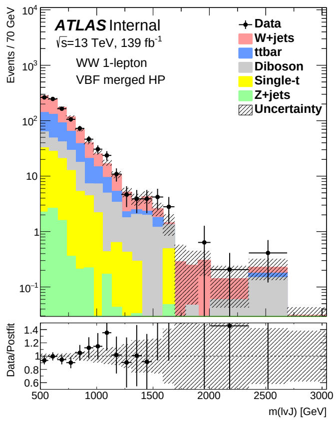
Channel	$V \to qq$ recon.	Signal regions	Background estimates: $W+jets$	$Z+jets$	$t\bar{t}$	Diboson	Single-t	Multijet	Total	Data
0-lepton (ZZ)			VBF category							
	Merged	HP	169 ± 12	228 ± 16	102 ± 10	51 ± 10	24 ± 4	-	574 ± 25	589
		LP	370 ± 23	411 ± 20	75 ± 8	30 ± 4	21 ± 4	-	906 ± 33	936
			ggF/DY category							
	Merged	HP Tag	133 ± 14	270 ± 40	437 ± 31	100 ± 10	45 ± 7	-	982 ± 60	978
		HP Untag	7600 ± 400	14 300 ± 600	6030 ± 270	2300 ± 180	840 ± 110	-	31 100 ± 800	31 074
		LP Tag	259 ± 28	560 ± 50	342 ± 24	67 ± 7	43 ± 7	-	1270 ± 70	1277
		LP Untag	16 300 ± 900	28 600 ± 1100	5040 ± 220	1760 ± 150	600 ± 80	-	52 400 ± 1500	52 396
1-lepton (WW)			VBF category							
	Merged	HP	530 ± 28	8.3 ± 0.5	321 ± 22	141 ± 27	113 ± 21	-	1110 ± 50	1096
		LP	1380 ± 40	24.5 ± 1.1	228 ± 17	150 ± 33	83 ± 16	-	1870 ± 60	1846
	Resolved		11 360 ± 190	530 ± 10	4060 ± 130	590 ± 80	1070 ± 210	960 ± 110	18 570 ± 340	18 530
			ggF/DY category							
	Merged	HP	24 820 ± 170	463 ± 5	13 890 ± 220	4910 ± 250	2800 ± 400	-	46 900 ± 500	47 330
		LP	60 270 ± 240	1095 ± 8	11 050 ± 160	3950 ± 210	1970 ± 250	-	78 300 ± 400	78 380
	Resolved		443 500 ± 1800	12 480 ± 40	126 000 ± 1500	16 800 ± 1200	21 200 ± 2800	27 200 ± 1400	647 000 ± 4000	645 610
2-lepton (ZZ)			VBF category							
	Merged	HP	0	87 ± 6	0.081 ± 0.009	9.6 ± 1.2	0	-	97 ± 6	101
		LP	0.133 ± 0.011	170 ± 8	0.85 ± 0.07	9.9 ± 1.2	0.43 ± 0.07	-	181 ± 8	162
	Resolved		0.272 ± 0.012	1566 ± 29	17.0 ± 0.7	72 ± 10	0.48 ± 0.32	-	1656 ± 31	1685
			ggF/DY category							
	Merged	HP Tag	0.0135 ± 0.0043	85 ± 6	0.283 ± 0.035	21.1 ± 2.3	0.34 ± 0.05	-	107 ± 7	94
		HP Untag	0.772 ± 0.010	3300 ± 40	4.27 ± 0.08	361 ± 32	0.58 ± 0.11	-	3670 ± 50	3 671
		LP Tag	0.0135 ± 0.0043	138 ± 8	0.313 ± 0.034	12.8 ± 1.4	0.30 ± 0.04	-	152 ± 8	141
		LP Untag	2.341 ± 0.017	5920 ± 50	10.16 ± 0.16	278 ± 26	2.03 ± 0.29	-	6220 ± 60	6095
	Resolved	Tag	0	1323 ± 26	110 ± 10	159 ± 12	4.7 ± 0.8	-	1600 ± 30	1583
		Untag	4.681 ± 0.026	42 750 ± 160	110.6 ± 1.5	1800 ± 100	13.4 ± 2.0	-	44 650 ± 190	44 604

Table 6.2: The expected background events with breakdowns from individual sources in 6 WW and 15 ZZ SRs compared with the data. The backgrounds are estimated from a background-only simultaneous fit to all WW and ZZ SRs and their corresponding CRs.

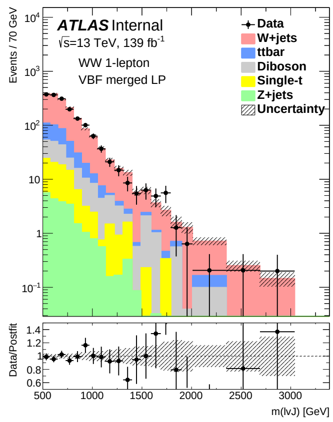


(a) VBF, merged HP
(b) VBF, merged LP
(c) ggF/DY, merged HP tagged
(d) ggF/DY, merged HP untagged
(e) ggF/DY, merged LP tagged
(f) ggF/DY, merged LP untagged

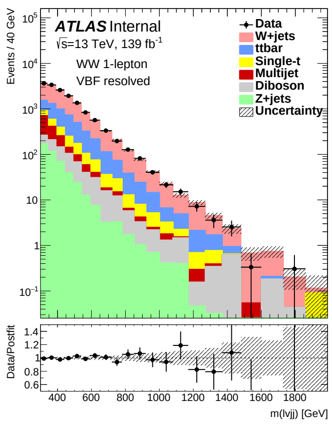
Figure 6.1: Comparisons of the observed data and the expected background distributions of m_T in the 6 ZZ SRs of the 0-lep channel. The background predictions are obtained through a background-only simultaneous fit to the 6 WW and 15 ZZ SRs and their respective $V + jets$ and $t\bar{t}$ CRs (see text). The middle panes show the ratios of the observed data to the background predictions. The bottom panes show the ratios of the post-fit and pre-fit background predictions. The hatched bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.



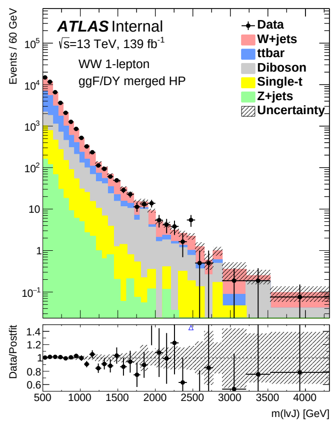
(a) VBF, merged HP



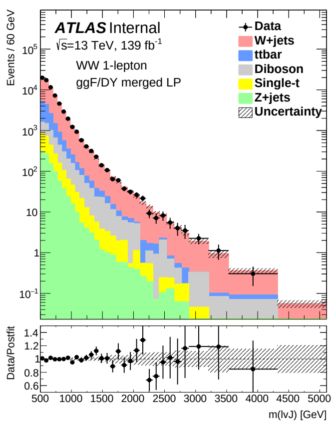
(b) VBF, merged LP



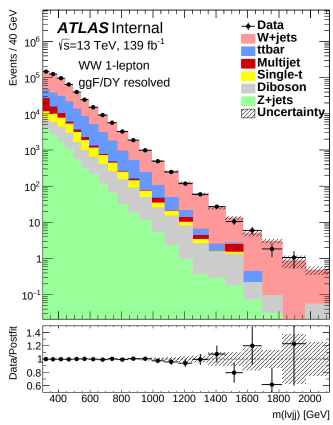
(c) VBF, resolved



(d) ggF/DY, merged HP

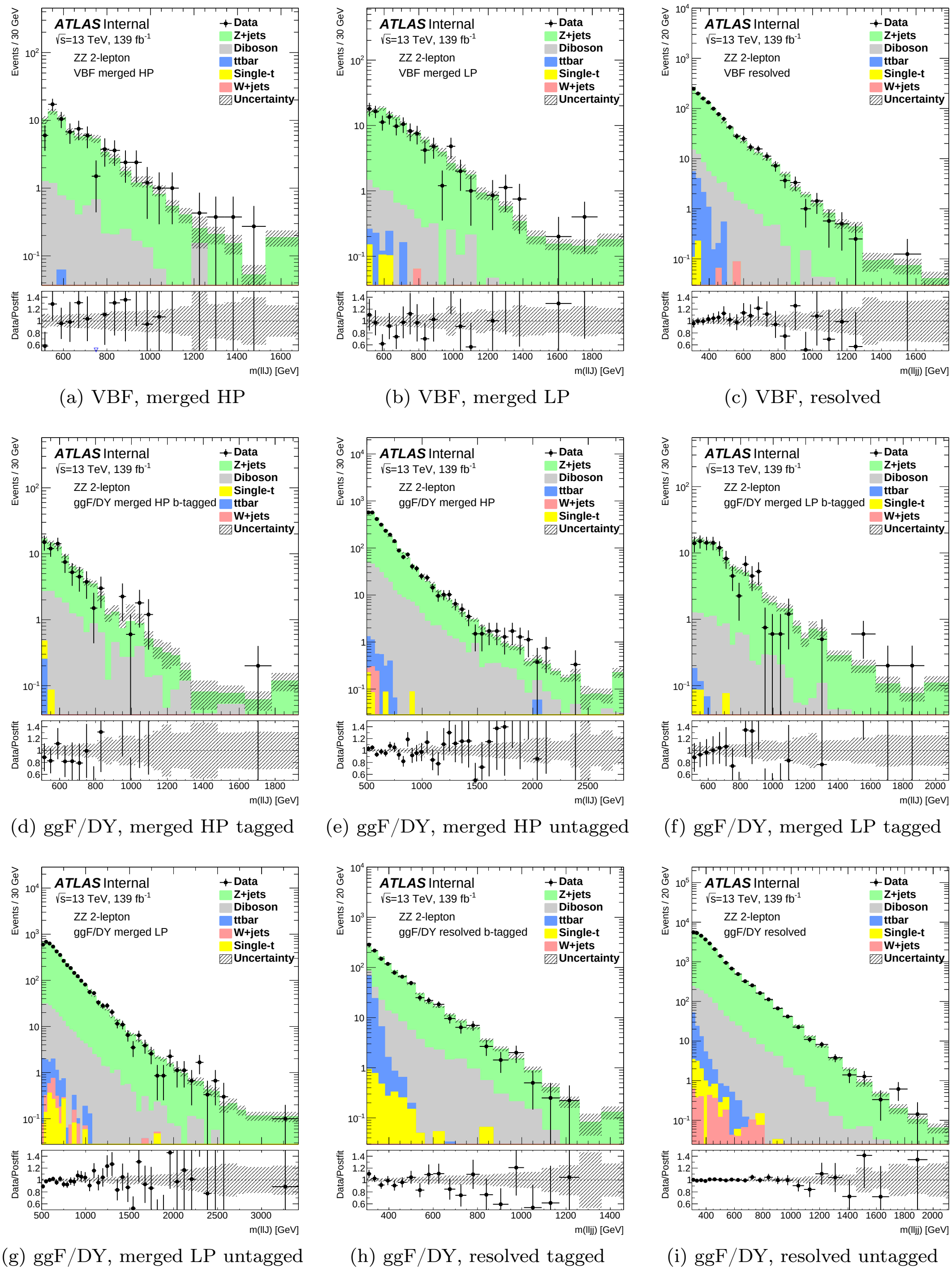


(e) ggF/DY, merged LP



(f) ggF/DY, resolved

Figure 6.2: Comparisons of the observed data and the expected background distributions of $m_{l\nu jj}$ or $m_{l\nu J}$ in the 6 WW SRs of the 1-lep channel. The background predictions are obtained through a background-only simultaneous fit to the 6 WW and 15 ZZ SRs and their respective $V + jets$ and $t\bar{t}$ CRs (see text). The middle panes show the ratios of the observed data to the background predictions. The bottom panes show the ratios of the post-fit and pre-fit background predictions. The hatched bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.

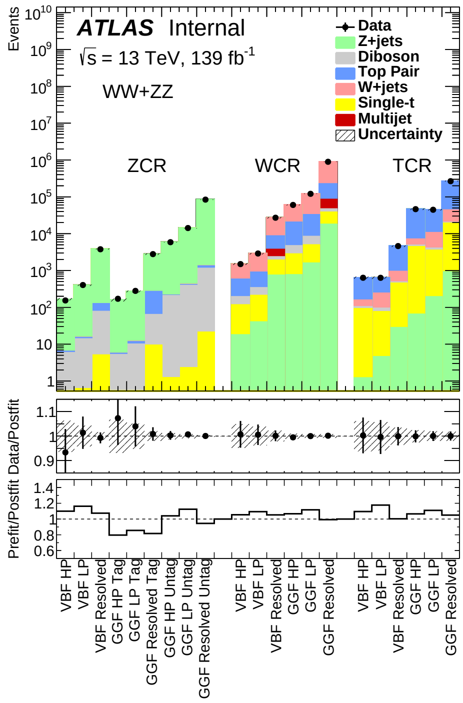


(a) VBF, merged HP (b) VBF, merged LP (c) VBF, resolved

(d) ggF/DY, merged HP tagged (e) ggF/DY, merged HP untagged (f) ggF/DY, merged LP tagged

(g) ggF/DY, merged LP untagged (h) ggF/DY, resolved tagged (i) ggF/DY, resolved untagged

Figure 6.3: Comparisons of the observed data and the expected background distributions of $m_{llj j}$ or m_{llJ} in the 9 ZZ SRs of the 2-lep channel. The background predictions are obtained through a background-only simultaneous fit to the 6 WW and 15 ZZ SRs and their respective $V + jets$ and $t\bar{t}$ CRs (see text). The middle panes show the ratios of the observed data to the background predictions. The bottom panes show the ratios of the post-fit and pre-fit background predictions. The hatched bands represent the uncertainties on the total background predictions, combining statistical and systematic contributions.



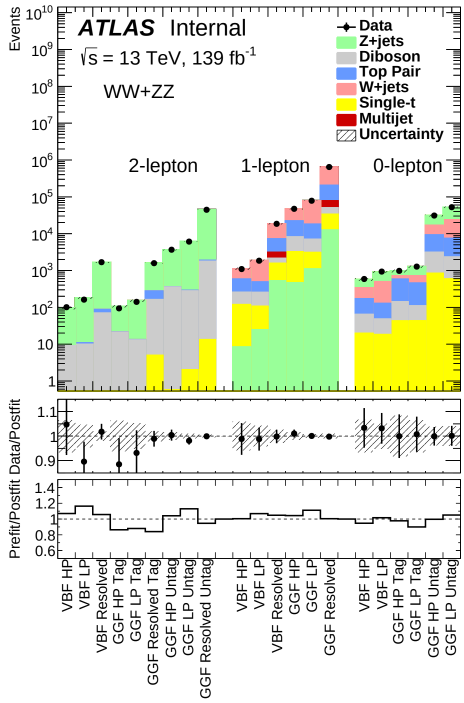


Figure 6.4: Comparisons of the observed data and the expected background event yields in (a) control and (b) signal regions for the $WW + ZZ$ fit. The background predictions are obtained through a background-only simultaneous fit to the 6 WW and 15 ZZ signal regions and their corresponding 21 $V + jets$ and $t\bar{t}$ CRs. The middle pane shows the ratios of the observed data to the post-fit background predictions. The bottom pane shows the ratios of the post-fit and pre-fit background predictions. The hatched bands represent uncertainties on the total background predictions, combining statistical and systematic contributions.

6.1.3 Limits on production of heavy resonances

In the absence of significant deviations from expected background contributions, constraints on the production cross sections of heavy resonances are derived by repeating the fit to the signal-plus-background hypothesis for different signal models. The exclusion limits are calculated with a modified frequentist method [147], also known as CLs, using the $\tilde{q_\mu}$ test statistic in the asymptotic approximation [148].

The upper limits on the production cross section of a Radion, $\sigma(pp \to R \to VV)$, are obtained by combining searches of $R \to WW$ and $R \to ZZ$ decays (assuming their branching ratios in Table 1.2) 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 $\sigma(pp \to R \to VV)$ upper limits at 95% confidence level (CL) as functions of its mass for both the ggF and VBF processes are shown in Figures 6.5. The observed (expected) limit varies from 1.8 (3.3) pb at 300 GeV to 0.44 (0.43) fb at 5000 GeV for the ggF production process and from 0.6 (1.15) pb at 300 GeV to 0.28 (0.27) fb at 5000 GeV for the VBF production process. The observed (expected) upper limits exclude the ggF production of a Radion with its mass below 2.9 (3.0) TeV while no mass exclusion can be made for the VBF production.

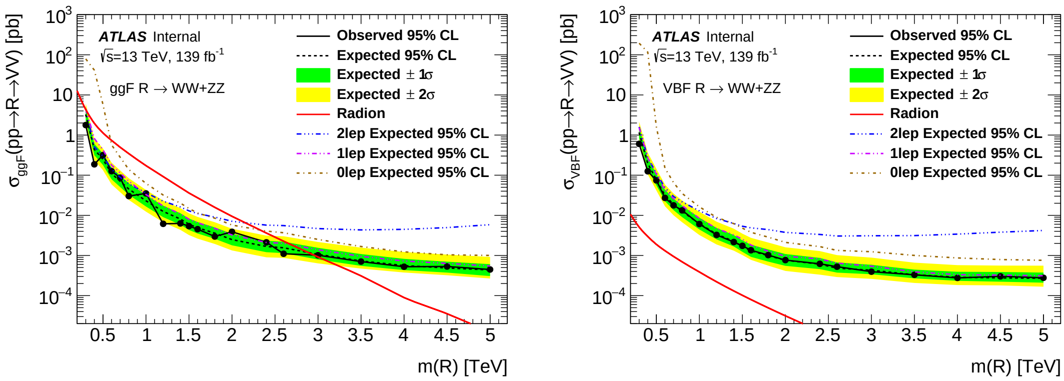


Figure 6.5: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at $\sqrt{s} = 13$ TeV of a Radion as functions of its mass for the (a) ggF and (b) VBF processes, combining searches for the WW and ZZ decays of the Radion in the three leptonic channels. The green (inner) and yellow (outer) bands represent $\pm1\sigma$ and $\pm2\sigma$ uncertainty in the expected limits. Limits expected from individual leptonic channels (dashed curves in purple, magenta, and blue) are also shown for comparison. Theoretical predictions as functions of the Radion mass are overlaid.

The upper limits on the production cross sections of HVT W' and Z' bosons are obtained through the WZ and $WW + ZZ$ fits, respectively. As shown in Table 6.1, all leptonic channels contribute to the $W' \to WZ$ search while only the 1-lep channel contributes to the $Z' \to WW$ search. Limits for W' and Z' production times their decay branching ratios to WZ or ZZ as functions of resonance masses are shown in Figures 6.6 and 6.7, respectively, for both DY and VBF processes. The theoretical predictions of the HVT Model A, Model B, and Model C are also shown for comparison. The observed (expected) $\sigma(pp \to W' \to WZ)$ limit range from 1.9 (2.5) pb at 300 GeV to 0.16 (0.17) fb at 5 TeV for DY production and from 1.3 (1.8) pb at 300 GeV to 0.35 (0.51) fb at 4 TeV for VBF production. These observed (expected) limits exclude an HVT W' boson produced in the DY process lighter than 3.9 (3.8) TeV for Model A and 4.3 (4.0) TeV for Model B, but fail to exclude any mass region in the VBF process.

Similarly, the observed (expected) $\sigma(pp \to Z' \to WW)$ limit range from 0.9 (2.7) pb at 300 GeV to 0.18 (0.18) fb at 5 TeV for the DY process and from 1.36 (3.15) pb at 300 GeV to 0.25 (0.32) fb at 4 TeV for the VBF process. These limits exclude an $HVTZ'$ boson lighter than 3.5 (3.4) TeV for Model A and 4.9 (3.7) TeV for Model B in the DY process.

The upper limits on the production cross section of the RSG graviton, $\sigma(pp \to G_{KK} \to VV)$,

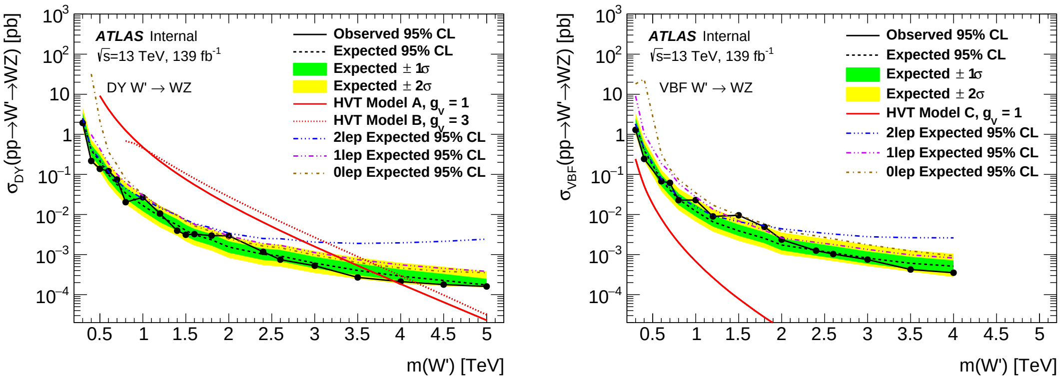


Figure 6.6: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at $\sqrt{s} = 13$ TeV of an HVT W' boson as functions of its mass for the (a) DY and (b) VBF processes, combining searches in the three leptonic channels for the $W' \to WZ$ decay. 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 purple, magenta, and blue) are also shown for comparison. Theoretical predictions as functions of the W' boson mass are overlaid in (a) for Model A and Model B and in (b) for Model C.

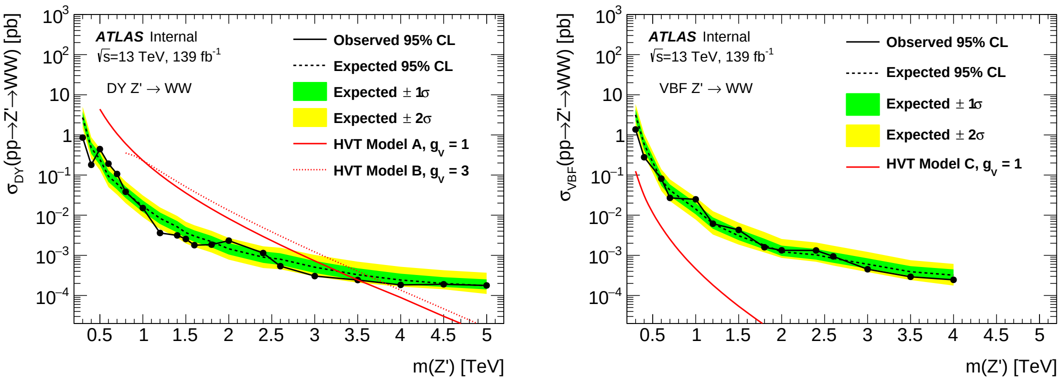


Figure 6.7: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at $\sqrt{s} = 13$ TeV of an HVT Z' boson as functions of its mass for the (a) DY and (b) VBF processes from the search for the $Z' \to WW$ decay in the 1-lep channel. The green (inner) and yellow (outer) bands represent $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty in the expected limits. Theoretical predictions as functions of the Z' boson mass are overlaid in (a) for Model A and Model B and in (b) for Model C.

are obtained following the same procedure used to derive the Radion limits. The observed and expected upper limits as functions of its mass for both the ggF and VBF processes are shown in Fig. 6.8. The observed (expected) limit varies from 1.4 (3.6) pb at 300 GeV to 0.26 (0.28) fb at 5 TeV for the ggF production process and from 0.40 (0.61) pb at 300 GeV to 0.30 (0.33) fb at 5 TeV for the VBF production process. Compared with theory cross sections, the observed (expected) upper limits exclude the production of a RS graviton at 95% CL lighter than 2.0 (2.2) TeV in the ggF process and lighter than 0.76 (0.77) TeV in the VBF process.

The effects of systematic uncertainties on the search are studied for hypothesised signals using the signal-strength parameter μ, the ratio between the extracted (Section **??**) and injected hypothesised signal cross sections. The expected relative uncertainties in the best-fit μ value from the leading sources of systematic uncertainty are shown in Table 6.3 for the ggF production of a RS graviton with $m(G_{KK}) = 600$ GeV and 2 TeV. Apart from the statistical uncertainties in the data, the uncertainties with the largest impact on the sensitivity of the searches are from the

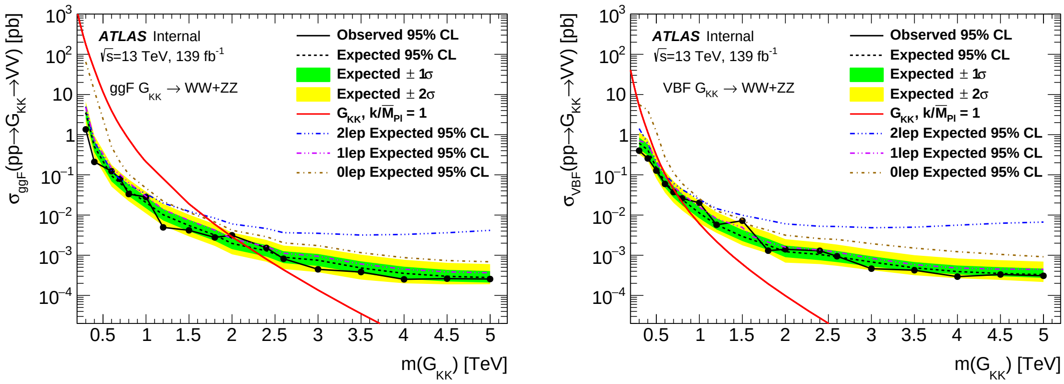


Figure 6.8: Observed (black solid curve) and expected (black dashed curve) 95% CL upper limits on the production cross section at $\sqrt{s} = 13$ TeV of a RSG graviton boson as functions of its mass for the (a) ggF and (b) VBF processes, combining searches for the $G_{KK} \to WW$ and $G_{KK} \to ZZ$ decay modes in the three leptonic channnels. The green (inner) and yellow (outer) bands represent $\pm1\sigma$ and $\pm2\sigma$ uncertainty in the expected limits. Limits expected from individual leptonic channels (dashed curves in purple, magenta, and blue) are also shown for comparison. Theoretical predictions for the chosen model as functions of the graviton boson mass are overlaid.

size of the MC samples, background normalisations, 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 searches are similar.

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

Table 6.3: Dominant relative uncertainties in the best-fit signal-strength parameter μ 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.

6.2 Results of the search for the non resonant electroweak diboson production

The statistical analysis is based on the profile likelihood test statistic [149] implemented with the RooFit [146] and RooStats [150] packages as similar to what done for the resonant statistical analysis.

A binned likelihood function $\mathcal{L}(\mu, \theta)$ is constructed as a product of Poisson probabilities over all of the bins of the distributions considered as template fit in the analysis.

This function depends on the signal-strength parameter μ, a multiplicative factor applied to the theoretical signal production cross section, and θ, a set of nuisance parameters that parametrise the effects of systematic uncertainties in the signal and expected backgrounds.

The binning is chosen so that the expected numbers of events ensure that the statistical uncertainty is less than 5% in most bins, while finer binning is also allowed in signal-enriched regions. The nuisance parameters are either free to float, or constrained using Gaussian or log-normal terms defined by external studies. The likelihood function for the combination of the three channels is the product of the Poisson likelihoods of the individual channels. However, only one constraint term per common nuisance parameter is included in the product.

A simultaneous maximum-likelihood fit is performed to the observed distributions of the final discriminants, BDT outputs, in the nine SRs to extract the signal rate information.

The three ZCRs, WCRs and TopCRs as well as the three VjjCRs are included in the fit's likelihood calculation; the m_{jj}^{Tag} distributions are used for ZCRs, WCRs and VjjCRs, while for the TopCRs only one bin for each of the three Vhad decay channels is used. The purpose of using m_{jj}^{Tag} distributions for CRs is to constrain the m_{jj}^{Tag} reweighting systematic uncertainties. The different regions and the corresponding discriminants entering the likelihood fit are summarized in Table 6.4. Signal and background contributions, including their shapes in the signal and control regions, are taken from MC simulations. For each source of systematic uncertainty, the correlations across bins of BDT distributions are taken into account and are fully correlated. The correlations between different regions, as well as those between signal and background, are also included. Moreover, normalization scale factors (SFs) are applied to the MC estimates of the Z+jets, W+jets and top quark contributions. These SFs are free parameters in the fit and are therefore constrained by the data in both the signal and control regions. The diboson contribution is constrained to the theoretical estimate within the corresponding uncertainties.

Regions		Discriminants		
		Merged high-purity	Merged low-purity	Resolved
0-leptons	SR	BDT	BDT	BDT
	VjjCR	m_{jj}^{Tag}	m_{jj}^{Tag}	m_{jj}^{Tag}
1-leptons	SR	BDT	BDT	BDT
	WCR	m_{jj}^{Tag}	m_{jj}^{Tag}	m_{jj}^{Tag}
	TopCR	One bin	One bin	One bin
2-leptons	SR	BDT	BDT	BDT
	ZCR	m_{jj}^{Tag}	m_{jj}^{Tag}	m_{jj}^{Tag}

Table 6.4: The distributions used in the global likelihood fit for the signal regions and control regions for all the categories in each channel. "One bin" implies that a single bin without any shape information is used in the corresponding fit region.

In general, one SF is introduced for each background component, common to both the SRs and CRs. One common Zjets SF is used for both the 0-leptons and 2-leptons channels, and one common Wjets SF is used for both the 0-leptons and 1-leptons channels. Similarly, one common ttbar SF is used for both the 0-leptons and 1-leptons channels. However, independent SFs are used for the resolved and merged categories, to take into account different MC modellings in the different phase spaces of the same background component.

The test statistic q_μ is defined as the profile likelihood ratio [151], $q_\mu = -2\ln\Lambda_\mu$ with $\Lambda_\mu = \mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)/\mathcal{L}(\hat{\mu}, \hat{\theta})$, where $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximize the likelihood function (with the constraint $0 \leq \hat{\mu} \leq \mu$), and $\hat{\hat{\theta}}_\mu$ are the values of the nuisance parameters that maximize the likelihood function for a given value of μ. The best-fit signal strength $\hat{\mu}$ value (muVBSobs) is obtained by maximizing the likelihood function with respect to all parameters. To determine whether the observed data is compatible with the background-only hypothesis, a test statistic $q_0 = -2\ln\Lambda_0$ is used.

Figures 6.9 and 6.10 show a selection of representative post-fit distributions of input variables that are most discriminating for each of the lepton channels, for the merged and resolved categories, respectively. Background and EW $VV+jj$ signal contributions shown are obtained from the signal-plus-background fits described previously.

The observed distributions of the BDT outputs in SRs used in the global likelihood fit are compared with the predictions, shown in Figure 6.11 for the 0 leptons channel, Figure 6.12 for the 1 lepton channel, and Figure 6.13 for the 2 leptons channel. The data distributions are reasonably well reproduced by the predicted contributions in all cases, with the smallest p-value of 0.16 from the χ^2 test [**chitest**] being for the m_{jj}^{Tag} distribution in the merged high-purity ZCR. The numbers of events observed and estimated in the SRs are summarized in Table 6.5 for the 0 leptons channel, Table 6.6 for the 1 lepton channel, and Table 6.7 for the 2 leptons channel. The fitted value of the signal strength is:

$$\mu^{\text{obs}}_{\text{EW}\,VVjj} = 1.05^{+0.42}_{-0.40} = 1.05 \pm 0.20(\text{stat.})^{+0.37}_{-0.34}(\text{syst.}). \tag{6.1}$$

The background-only hypothesis is excluded in data with a significance of 2.7 standard deviations, compared with 2.5 standard deviations expected.

Figure 6.14 shows the measured signal strength from the combined fit with a single signal-strength fit parameter, and from a fit where each lepton channel has its own signal-strength parameter. The probability that the signal strengths measured in the three lepton channels are compatible is 36%.

Sample		Resolved	Merged HP	Merged LP
Background	W + jets	9200 ± 1300	259 ± 27	582 ± 56
	Z + jets	$19\,000 \pm 1400$	383 ± 29	955 ± 69
	Top quarks	3280 ± 480	277 ± 28	276 ± 32
	Diboson	720 ± 120	69 ± 12	68 ± 14
	Total	$32\,100 \pm 2000$	988 ± 50	1881 ± 96
Signal	$W(\ell\nu)W(qq')$	56 ± 22	8.0 ± 3.2	5.4 ± 2.2
	$W(\ell\nu)Z(qq)$	12.0 ± 4.7	2.1 ± 0.8	1.6 ± 0.6
	$Z(\nu\nu)W(qq')$	66 ± 25	9.0 ± 3.5	7.4 ± 2.9
	$Z(\nu\nu)Z(qq)$	27 ± 10	5.1 ± 2.0	3.1 ± 1.2
	Total	161 ± 35	24.3 ± 5.2	17.5 ± 3.9
SM		$32\,300 \pm 2000$	1012 ± 50	1898 ± 96
Data		32 299	1002	1935

Table 6.5: Numbers of events observed and predicted for signal and background processes in the 0-lepton channel signal regions, obtained from signal-plus-background fits to the signal and control regions. The signal yields are calculated after the fit with the observed signal strength of 1.05 applied. The uncertainties combine statistical and systematic contributions. The fit constrains the background estimate towards the observed data, which reduces the total background uncertainty by correlating those uncertainties from the individual backgrounds.

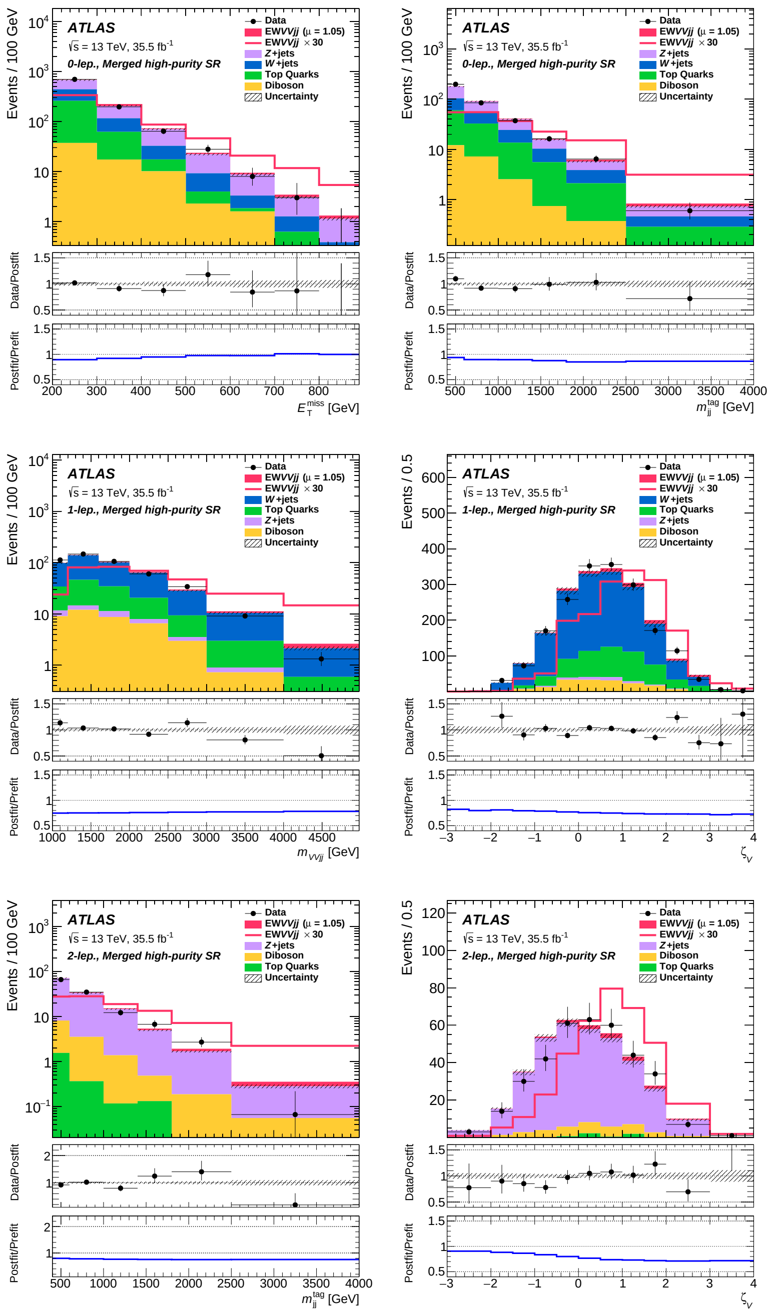


Figure 6.9: The distributions for E_T^{miss} (top left), m_{jj}^{Tag} (top right), m_{VVjj} (middle left), ζ_V (middle right), m_{jj}^{Tag} (bottom left), and ζ_V (bottom right) in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channels for the high-purity merged signal region. The background contributions after the global likelihood fit are shown as filled histograms. The signal is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.05$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The middle pane shows the ratios of the observed data to the post-fit signal and background predictions. The bottom pane shows the ratios of the post-fit and pre-fit background predictions.

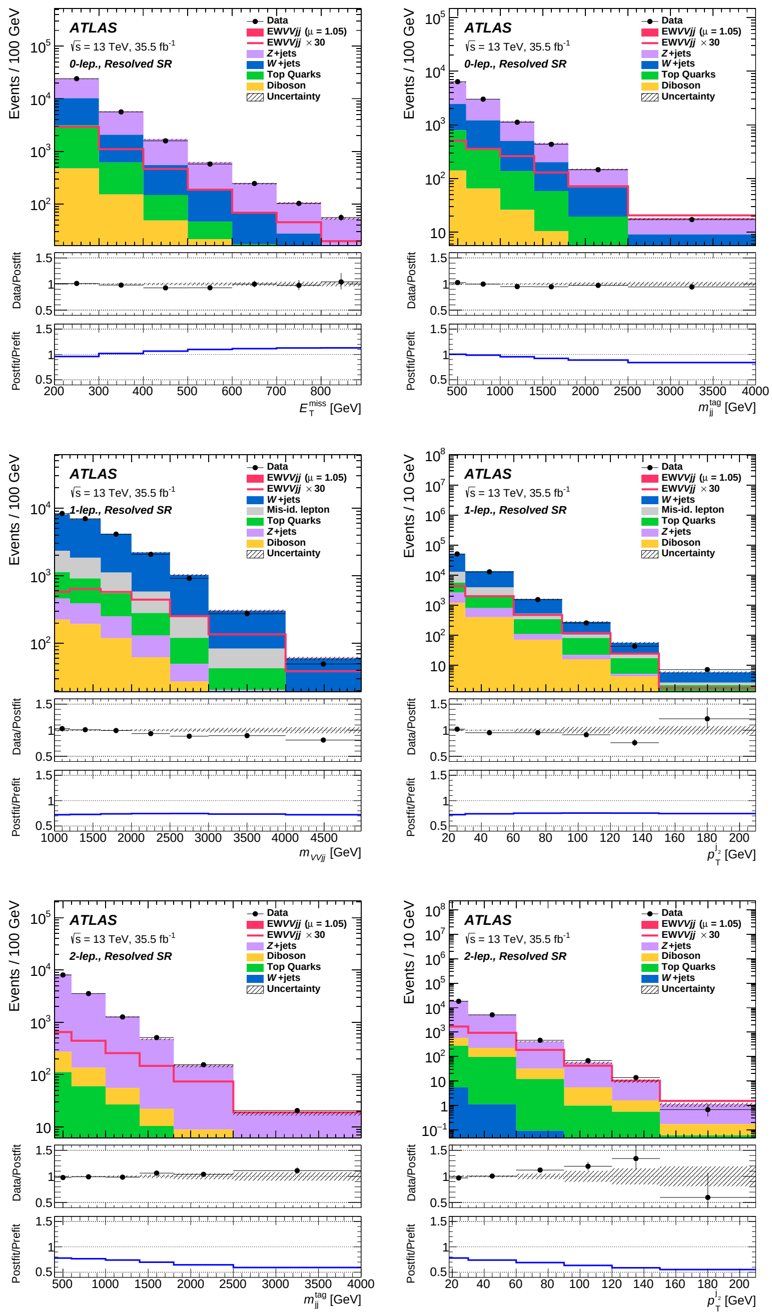


Figure 6.10: The distributions for E_T^{miss} (top left), m_{jj}^{Tag} (top right), m_{VVjj} (middle left), $p_{T,j2}$ (middle right), m_{jj}^{Tag} (bottom left), and $p_{T,j2}$ (bottom right) in the 0-lepton (top), 1-lepton (middle) and 2-lepton (bottom) channels for the resolved signal region. The background contributions after the global likelihood fit are shown as filled histograms. The signal is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.05$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The middle pane shows the ratios of the observed data to the post-fit signal and background predictions. The bottom pane shows the ratios of the post-fit and pre-fit background predictions.

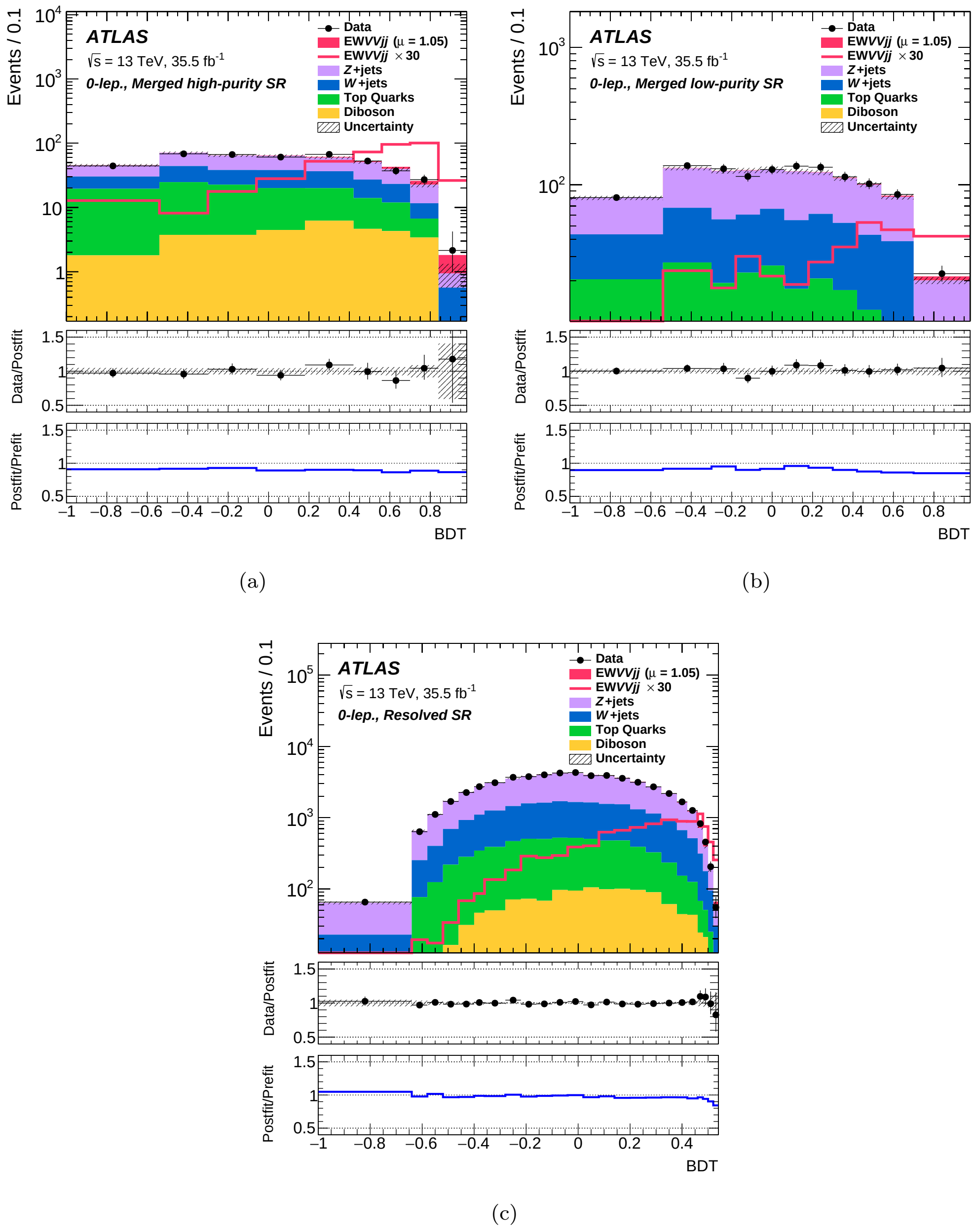


Figure 6.11: Comparisons of the observed data and expected distributions of the BDT outputs of the 0-lepton channel signal regions: (a) high-purity and (b) low-purity merged signal regions; (c) the resolved signal region. The background contributions after the global likelihood fit are shown as filled histograms. The signal is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.05$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The middle pane shows the ratios of the observed data to the post-fit signal and background predictions. The uncertainty in the total prediction, shown as bands, combines statistical and systematic contributions. The bottom pane shows the ratios of the post-fit and pre-fit background predictions.

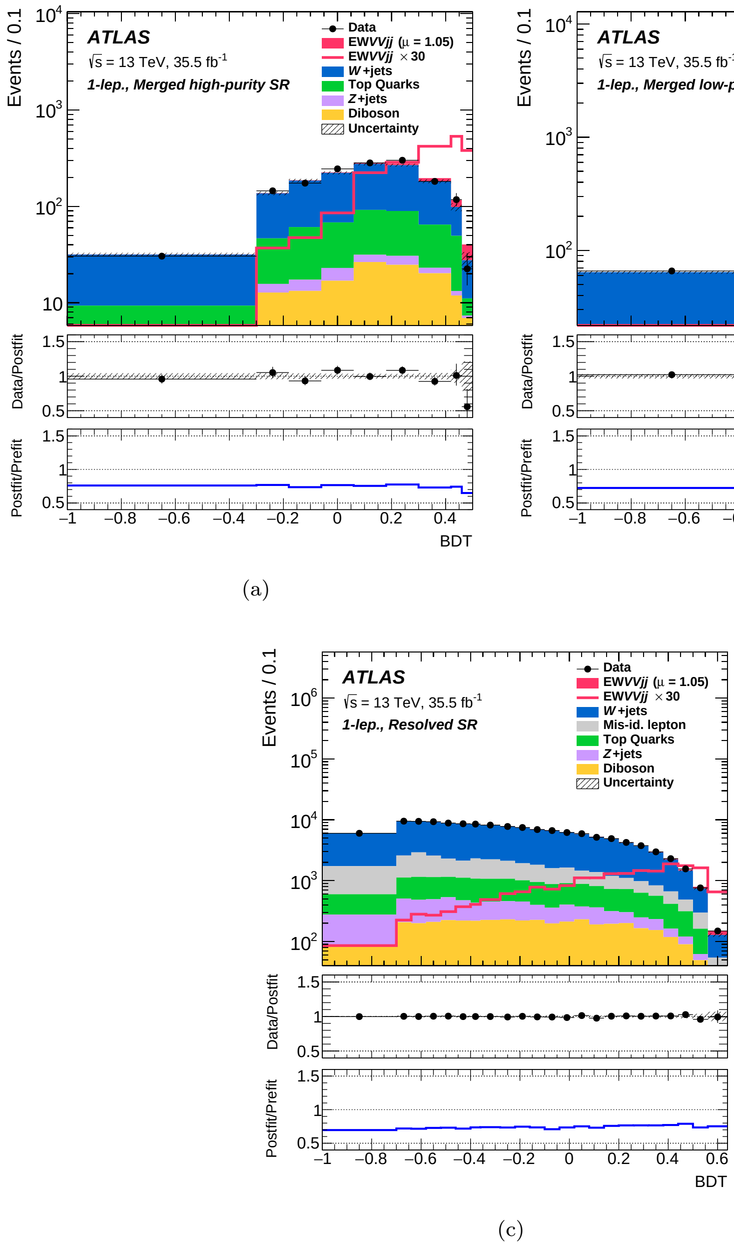
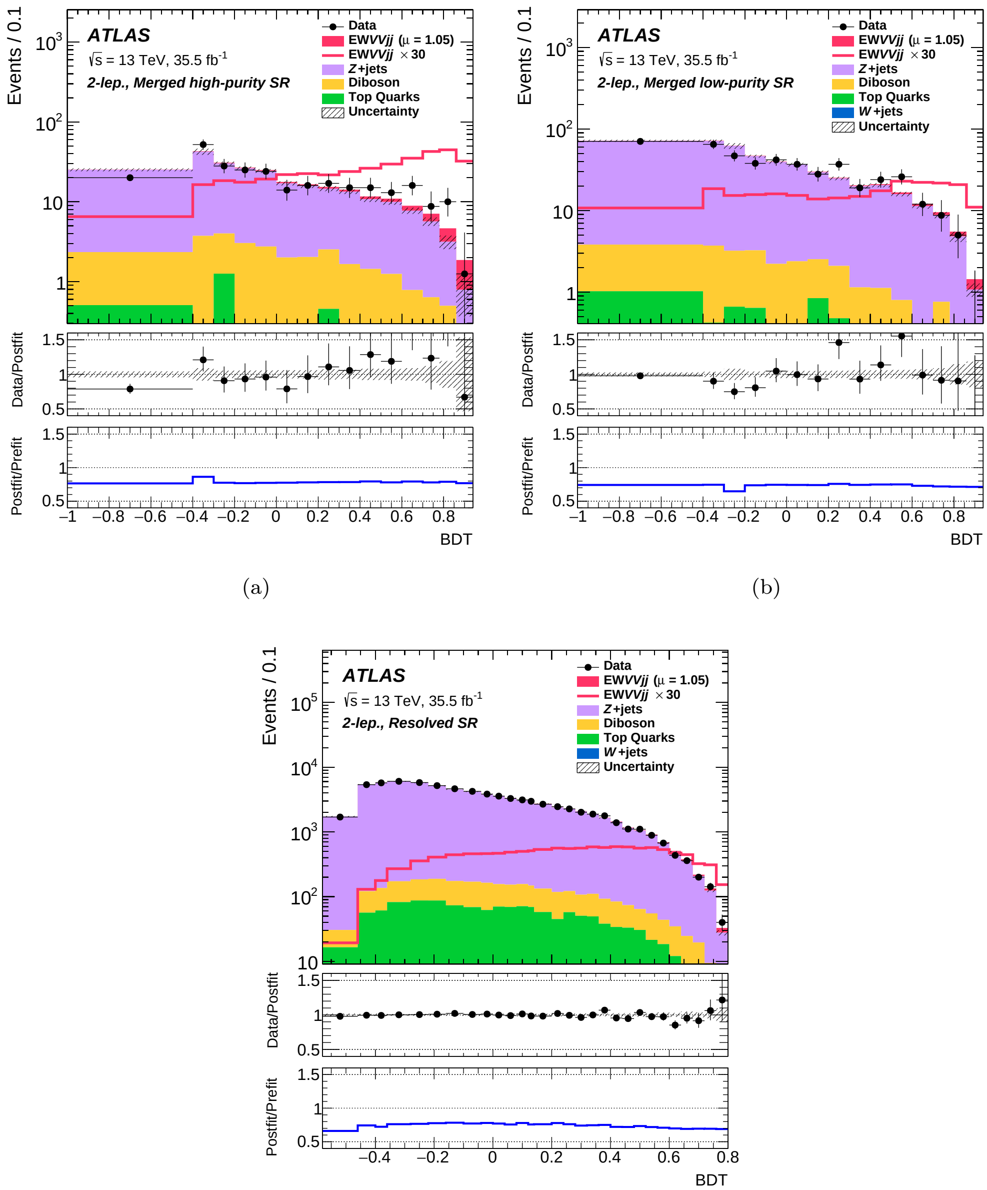


Figure 6.12: Comparisons of the observed data and expected distributions of the BDT outputs of the 1-lepton channel signal regions: (a) high-purity and (b) low-purity merged signal regions; (c) the resolved signal region. The background contributions after the global likelihood fit are shown as filled histograms. The signal is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.05$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The middle pane shows the ratios of the observed data to the post-fit signal and background predictions. The uncertainty in the total prediction, shown as bands, combines statistical and systematic contributions. The bottom pane shows the ratios of the post-fit and pre-fit background predictions.



(c)

Figure 6.13: Comparisons of the observed data and expected distributions of the BDT outputs of the 2-lepton channel signal regions: (a) high-purity and (b) low-purity merged signal regions; (c) the resolved signal region. The background contributions after the global likelihood fit are shown as filled histograms. The signal is shown as a filled histogram on top of the fitted backgrounds normalized to the signal yield extracted from data ($\mu = 1.05$), and unstacked as an unfilled histogram, scaled by the factor indicated in the legend. The entries in overflow are included in the last bin. The middle pane shows the ratios of the observed data to the post-fit signal and background predictions. The uncertainty in the total prediction, shown as bands, combines statistical and systematic contributions. The bottom pane shows the ratios of the post-fit and pre-fit background predictions.

Sample		Resolved	Merged HP	Merged LP
Background	W + jets	$69\,100 \pm 1900$	1201 ± 65	2828 ± 97
	Z + jets	2770 ± 370	39 ± 3	83 ± 6
	Top quarks	7100 ± 1100	394 ± 56	422 ± 63
	Diboson	2660 ± 600	163 ± 35	229 ± 57
	Multijet	$13\,400 \pm 1600$	–	–
	Total	$95\,100 \pm 2800$	1797 ± 93	3560 ± 130
Signal	$W(\ell\nu)W(qq')$	330 ± 120	45 ± 17	34 ± 13
	$W(\ell\nu)Z(qq)$	78 ± 29	11 ± 4	5 ± 2
	Total	410 ± 130	57 ± 18	39 ± 13
SM		$95\,500 \pm 2800$	1854 ± 95	3600 ± 130
Data		95 366	1864	3571

Table 6.6: Numbers of events observed and predicted for signal and background processes in the 1-lepton channel signal regions, obtained from signal-plus-background fits to the signal and control regions. The signal yields are calculated after the fit with the observed signal strength of 1.05 applied. The uncertainties combine statistical and systematic contributions. The fit constrains the background estimate towards the observed data, which reduces the total background uncertainty by correlating those uncertainties from the individual backgrounds.

Sample		Resolved	Merged HP	Merged LP
Background	Z + jets	$37\,090 \pm 310$	331 ± 14	775 ± 24
	Top quarks	645 ± 99	5.8 ± 0.9	9.9 ± 2.7
	Diboson	830 ± 170	34.6 ± 7.6	36.7 ± 8.2
	Total	$38\,570 \pm 370$	371 ± 16	821 ± 25
Signal	$Z(\ell\ell)W(qq')$	138 ± 53	8.6 ± 3.3	7.0 ± 2.7
	$Z(\ell\ell)Z(qq)$	46 ± 18	4.3 ± 1.7	2.9 ± 1.1
	Total	185 ± 56	12.9 ± 3.7	9.8 ± 2.9
SM		$38\,760 \pm 370$	384 ± 17	831 ± 25
Data		38 734	371	810

Table 6.7: Numbers of events observed and predicted for signal and background processes in the 2-lepton channel signal regions, obtained from signal-plus-background fits to the signal and control regions. The signal yields are calculated after the fit with the observed signal strength of 1.05 applied. The uncertainties combine statistical and systematic contributions. The fit constrains the background estimate towards the observed data, which reduces the total background uncertainty by correlating those uncertainties from the individual backgrounds.

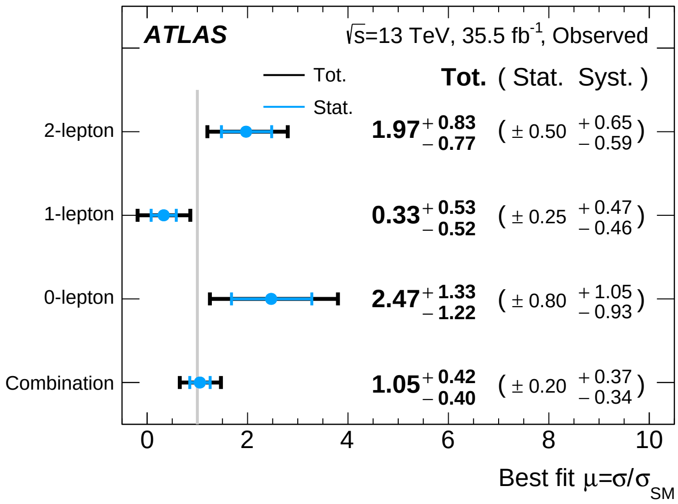


Figure 6.14: The fitted values of the signal-strength parameter $\mu^{\rm obs}_{{\rm EW}\,VVjj}$ for the 0-, 1- and 2-lepton channels and their combination. The individual $\mu^{\rm obs}_{{\rm EW}\,VVjj}$ values for the lepton channels are obtained from a simultaneous fit with the signal-strength parameter for each of the lepton channels floating independently. The probability that the signal strengths measured in the three lepton channels are compatible is 36%.

After the global maximum-likelihood fit, the uncertainties described in Section 5.10 are much reduced. The effects of systematic uncertainties on the measurement after the fit are studied using the signal-strength parameter $\mu^{\mathrm{obs}}_{\mathrm{EW}\,VVjj}$. The relative uncertainties in the best-fit $\mu^{\mathrm{obs}}_{\mathrm{EW}\,VVjj}$ value from the leading sources of systematic uncertainty are shown in Table 6.8. The individual sources of systematic uncertainty detailed in Section 5.10 are combined into categories. Apart from the statistics of the data, the uncertainties with the largest impact on the sensitivity of EW $VV+jj$ production are from the modeling of background ($Z+jets$, $W+jets$ and QCD-induced diboson processes), the modeling of the signal, b-tagging, and reconstruction of small-R and large-R jets.

Uncertainty source	σ_μ
Total uncertainty	0.41
Statistical	0.20
Systematic	0.35
Theoretical and modeling uncertainties	
Floating normalizations	0.09
Z+jets	0.13
W+jets	0.09
$t\bar{t}$	0.06
Diboson	0.09
Multijet	0.04
Signal	0.07
MC statistics	0.17
Experimental uncertainties	
Large-R jets	0.08
Small-R jets	0.06
Leptons	0.02
E_T^{miss}	0.04
b-tagging	0.07
Pileup	0.04
Luminosity	0.03

Table 6.8: The symmetrized uncertainty σ_μ from each source in the best-fit signal-strength parameter $\mu^{\mathrm{obs}}_{\mathrm{EW}\,VVjj}$. The floating normalizations include uncertainties of normalization scale factors for Z+jets, W+jets and top quark contributions.

6.2.1 Cross section measurements

The determination of the fiducial cross section is performed by scaling the measured signal strengths with the corresponding SM predicted fiducial cross sections, $\sigma^{\text{fid,obs}}_{\text{EW}\,VVjj} = \mu^{\text{obs}}_{\text{EW}\,VVjj} \cdot \sigma^{\text{fid,SM}}_{\text{EW}\,VVjj}$.

It is assumed that there is no new physics that could cause significant kinematic differences of the background and signal. Therefore, the only new physics signals that can be detected in an unbiased way are those leading to an enhanced EW $VV + jj$ signal strength in the search region of this analysis. The fiducial cross sections for EW $VV + jj$ are measured in the merged and resolved fiducial phase-space regions described in Section 5.13 and inclusively. The merged HP SR and LP SR are combined to form one single merged fiducial phase-space region. The systematic uncertainties of the measured fiducial cross sections include contributions from experimental systematic uncertainties, theory modeling uncertainties in the backgrounds, theory modeling uncertainties in the shapes of signal kinematic distributions, and luminosity uncertainties. The measured and SM predicted fiducial cross sections for EW $VV + jj$ processes are summarized in Tabl 6.9, where the measured values are obtained from two different simultaneous fits. In the first fit, two signal-strength parameters are used, one for the merged category (both HP and LP), and the other one for the resolved category; while in the second fit, a single signal-strength parameter is used. The measured and SM predicted fiducial cross sections in each lepton channel are also reported in Table 6.10. The measured values are obtained from a simultaneous fit where each lepton channel has its own signal-strength parameter, and in each lepton channel the same signal-strength parameter is applied to both the merged and resolved categories. The predictions are from $MadGraph5_aMC@NLO$ 2.4.3 at LO only, and no higher order corrections are included; the theoretical uncertainties due to the PDF, missing higher-order corrections, and parton-shower modeling are estimated as described in Section 5.10. The measured fiducial cross sections are generally consistent with the SM predictions.

Fiducial phase space	Predicted $\sigma^{\text{fid,SM}}_{\text{EW}\,VVjj}$ [fb]	Measured $\sigma^{\text{fid,obs}}_{\text{EW}\,VVjj}$ [fb]
Merged	11.4 ± 0.7 (theo.)	12.7 ± 3.8 (stat.) $^{+4.8}_{-4.2}$ (syst.)
Resolved	31.6 ± 1.8 (theo.)	26.5 ± 8.2 (stat.) $^{+17.4}_{-17.1}$ (syst.)
Inclusive	43.0 ± 2.4 (theo.)	45.1 ± 8.6 (stat.) $^{+15.9}_{-14.6}$ (syst.)

Table 6.9: Summary of predicted and measured fiducial cross sections for EW $VV+jj$ production. The three lepton channels are combined. For the measured fiducial cross sections in the merged and resolved categories, two signal-strength parameters are used in the combined fit, one for the merged category and the other one for the resolved category; while for the measured fiducial cross section in the inclusive fiducial phase space, a single signal-strength parameter is used. For the SM predicted cross section, the error is the theoretical uncertainty (theo.). For the measured cross section, the first error is the statistical uncertainty (stat.), and the second error is the systematic uncertainty (syst.).

Fiducial phase space		Predicted $\sigma^{\text{fid,SM}}_{\text{EW}\,VVjj}$ [fb]	Measured $\sigma^{\text{fid,obs}}_{\text{EW}\,VVjj}$ [fb]
Merged	0-lepton	$4.1 \pm 0.3\,(\text{theo.})$	$10.1 \pm 3.3\,(\text{stat.})\ ^{+4.2}_{-3.8}\,(\text{syst.})$
	1-lepton	$6.1 \pm 0.5\,(\text{theo.})$	$2.0 \pm 1.5\,(\text{stat.})\ ^{+2.9}_{-2.8}\,(\text{syst.})$
	2-lepton	$1.2 \pm 0.1\,(\text{theo.})$	$2.4 \pm 0.6\,(\text{stat.})\ ^{+0.8}_{-0.7}\,(\text{syst.})$
Resolved	0-lepton	$9.2 \pm 0.6\,(\text{theo.})$	$22.8 \pm 7.4\,(\text{stat.})\ ^{+9.4}_{-8.5}\,(\text{syst.})$
	1-lepton	$16.4 \pm 1.0\,(\text{theo.})$	$5.5 \pm 4.1\,(\text{stat.})\ ^{+7.7}_{-7.5}\,(\text{syst.})$
	2-lepton	$6.0 \pm 0.4\,(\text{theo.})$	$11.8 \pm 3.0\,(\text{stat.})\ ^{+3.8}_{-3.5}\,(\text{syst.})$
Inclusive	0-lepton	$13.3 \pm 0.8\,(\text{theo.})$	$32.9 \pm 10.7\,(\text{stat.})\ ^{+13.5}_{-12.3}\,(\text{syst.})$
	1-lepton	$22.5 \pm 1.5\,(\text{theo.})$	$7.5 \pm 5.6\,(\text{stat.})\ ^{+10.5}_{-10.2}\,(\text{syst.})$
	2-lepton	$7.2 \pm 0.4\,(\text{theo.})$	$14.2 \pm 3.6\,(\text{stat.})\ ^{+4.6}_{-4.2}\,(\text{syst.})$

Table 6.10: Summary of predicted and measured fiducial cross sections for EW $VVjj$ production in the three lepton channels. For the SM predicted cross section, the error is the theoretical uncertainty (theo.). For the measured cross section, the first error is the statistical uncertainty (stat.), and the second error is the systematic uncertainty (syst.).

Conclusions

In this Thesis work a detailed study on dibosonic events (*ZZ*, *ZW* or *WW*) in semi-leptonic decay channels have been performed. These final states could be used for searches of new resonances and for measurements of the cross-section in the non-resonant interpretation. A Machine Learning approach have been developed as powerful tools for the identification and the selection of the production mechanism of the new resonances predicted in BSM models, such as *Vector Boson Fusion* mechanism or the *Vector Boson Scattering* in the non-resonant channel.

Machine Learning techniques help in the improvement of the sensitivity and represents a challenging development of new techniques that could perform better and better in high energy experiments. For the first time in ATLAS a *Recurrent Neural Network* is used to perform the identification of the VBF topology in the production of new physics resonances; besides the trial of new Machine Learning techniques, this represented a concrete improvement in the approach of the classification problem of this topology. This new approach relies on the informations coming from all the jets reconstructed in the event. The RNN architecture is able to manage variable length sequences of inputs in order to maximise the informations used and exploiting the correlation in the sequences of the jets. Moreover, the VBF jets are not selected, so also events in which only one VBF jet is reconstructed are automatically recovered. The VBF signal efficiency has improved by a factor $10 - 60\%$ depending on the mass of the resonance.

The searches for the production of heavy resonances are performed in the diboson semi-leptonic channels using the full run-II dataset corresponding to an integrated luminosity of 139 fb^{-1}, recorded by ATLAS at $\sqrt{s} = 13\ TeV$. The data observed in the 3 lepton channels considered are found to be in good agreement with the background expectations. In the absence of significant excesses, upper limits on the production of heavy resonances in the mass range 300–5000 GeV through gluon-gluon fusion, Drell-Yan or vector-boson fusion processes are derived for SM extensions with an additional neutral scalar, a heavy vector triplet or warped extra dimensions. Combining the WW and ZZ decay modes, the observed 95% CL upper limit on $\sigma(pp \to X \to VV)$ for the ggF (VBF) process ranges from 1.8 (0.6) pb at 300 GeV to 0.44 (0.28) fb at 5 TeV for a Radion and from 1.4 (0.4) pb at 300 GeV to 0.26 (0.30) fb at 5 TeV for a RS graviton. These observed limits set lower mass limits of 2.9 TeV for the ggF production of a Radion, 2.0 (0.76) TeV for the ggF (VBF) production of a RS graviton.

For the production of W' and Z' bosons in the HVT framework, the observed upper limit on $\sigma(pp \to W' \to WZ)$ varies from 1.9 pb at 300 GeV to 0.16 fb at 5 TeV for DY production and from 1.3 pb at 300 GeV to 0.35 fb at 4 TeV for VBF production. Similarly, the limits on $\sigma(pp \to Z' \to WW)$ are observed to vary from 0.9 pb at 300 GeV to 0.18 fb at 5 TeV for DY production and from 1.36 pb at 300 GeV to 0.25 fb at 4 TeV for VBF production. In the benchmark Model A (Model B), these cross section limits exclude the DY production of a W' boson with $m(W') < 3.9$ (4.3) TeV and a Z' boson with $m(Z') < 3.5$ (3.9) TeV.

Moreover, a measurement of diboson electroweak production with association of two energetic jets are performed. The data were collected in 2015 and 2016 and correspond to a total integrated luminosity of 35.5 fb^{-1}. The $VVjj$ electroweak production cross section is measured with a significance of 2.7 standard deviations over the background-only hypothesis. The expected significance is 2.5 standard deviations. The measured signal strength relative to the leading-order SM prediction is $\mu^{obs}_{EWVVjj} = 1.05 \pm 0.20(stat.)^{+0.37}_{-0.34}(syst.)$. The fiducial cross section of $VVjj$ electroweak production is measured to be $\sigma^{fid,obs}_{EWVVjj} = 45.1 \pm 8.6(stat.)^{+15.9}_{-14.6}(syst.)\ fb$.

Appendix A

Appendix: Single Lepton Triggers in ATLAS

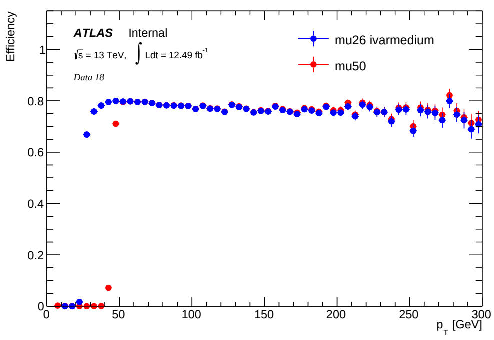
The lepton trigger efficiency have been evaluated using the Tag & Probe method using events with two same flavour leptons and passing the baseline Z selection described in Section 5.4.1.

The tag lepton is required to have a $p_T > 30\ GeV$. Data taken in the 2018 with a luminosity of 12.49 fb_{-1} have been used to study the trigger efficiency in the 2018. The trigger thresholds studied for muons are: HLT_mu26_ivarmedium and HLT_mu50. The trigger thresholds studied for electrons are: HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0 and HLT_e140_lhloose_nod0.

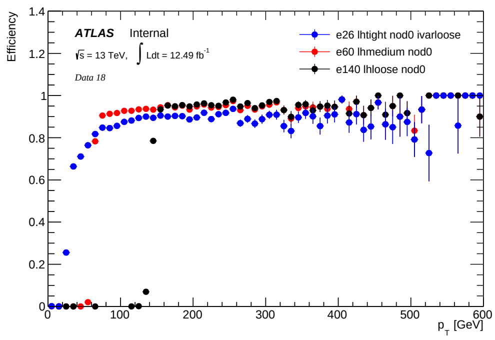
In Figure A.1 the efficiency as a function on the p_T (a) for muons and (b) for electrons for different trigger thresholds are shown.

In Figure A.2 efficiency as a function on the η (a) for muons and (b) for electrons for different trigger thresholds at their plateau is showing.

In Figure A.3 efficiency as a function on the average number of interactions per bunch crossing ($< \mu >$) (a) for muons and (b) for electrons for different trigger thresholds at their plateau is showed. We also performed a comparison of the efficiency curves using 2018 and 2017 data.



(a)



(b)

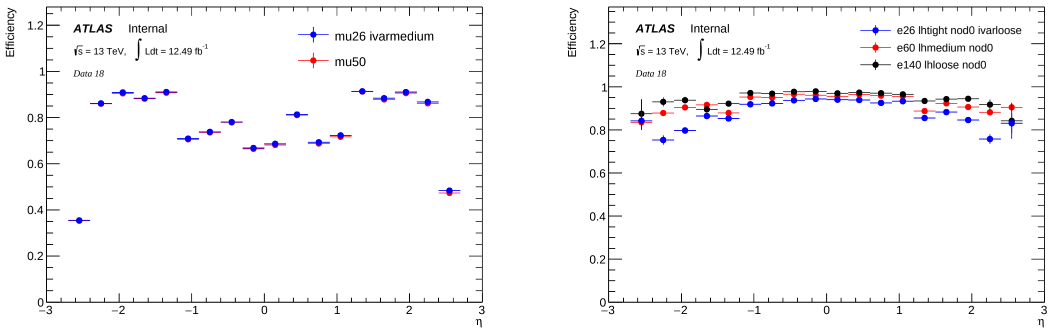
Figure A.1: Efficiency as a function on the p_T (a) for muons and (b) for electrons for different trigger thresholds.

For the 2017 a luminosity of 43.81 fb^{-1} have been used. In Figure A.4 the efficiency curves as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_mu26_ivarmedium are shown. The efficiency in η and $< \mu >$ are always taken at the trigger threshold at plateau.

In Figure A.5 the efficiency curves as a function of the p_T (a), η (b) and of the average number

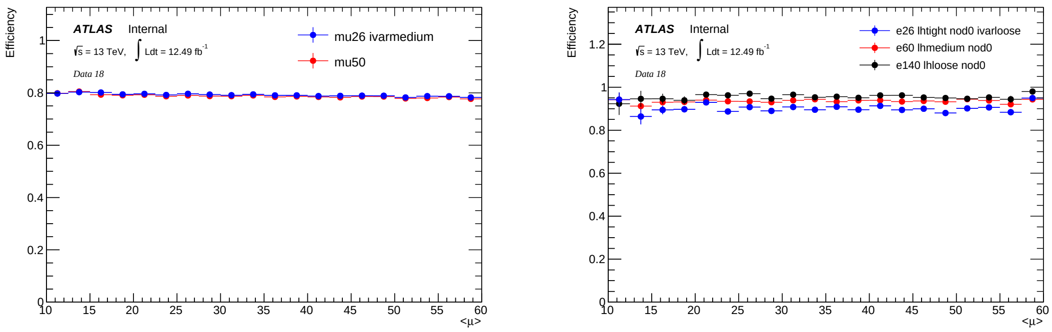
Data-taking period	$e\nu qq$ and $eeqq$ channels	$\mu\nu qq$ ($p_T(\mu\nu) < 150\ GeV$) and $\mu\mu qq$ channels	$\mu\nu qq$ ($p_T(\mu\nu) > 150\ GeV$) and $\nu\nu qq$ channels
2015	HLT_e24_lhmedium_L1EM20 OR HLT_e60_lhmedium OR HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 OR HLT_mu50	HLT_xe70
2016a (run < 302919) ($L < 1.0 \times 10^{34}\,cm^{-2}\,\mathrm{s}^{-1}$)	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e140_lhloose_nod0 HLT_e300_etcut	HLT_mu26_ivarmedium OR HLT_mu50	HLT_xe90_mht_L1XE50
2016b (run ≥ 302919) ($L < 1.7 \times 10^{34}\,cm^{-2}\,\mathrm{s}^{-1}$)	same as above	same as above	HLT_xe110_mht_L1XE50
2017	same as above	same as above	HLT_xe110_pufit_L1XE55
2018	same as above	same as above	HLT_xe110_pufit_xe70_L1XE50

Table A.1: The list of triggers used in the analysis.



(a) (b)

Figure A.2: Efficiency as a function on the η (a) for muons and (b) for electrons for different trigger thresholds at their plateau.



(a) (b)

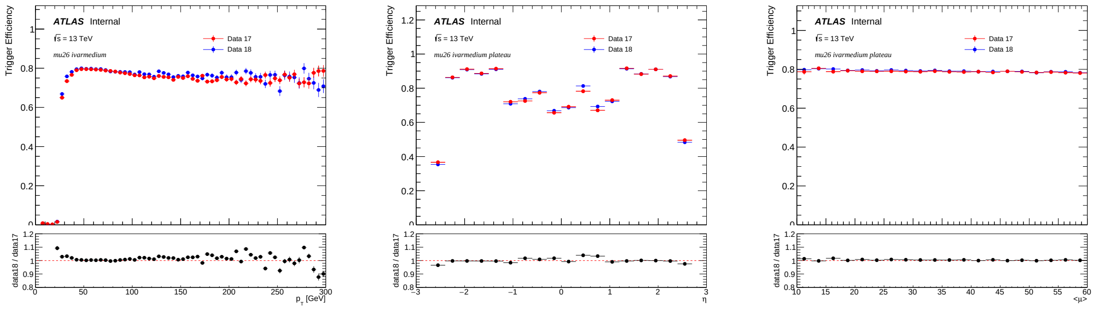
Figure A.3: Efficiency as a function on the average number of interactions per bunch crossing ($< \mu >$) (a) for muons and (b) for electrons for different trigger thresholds at their plateau.

of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_mu50 are shown.

In Figure A.6 the efficiency curves as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_e26_lhtight_nod0_ivarloose are shown.

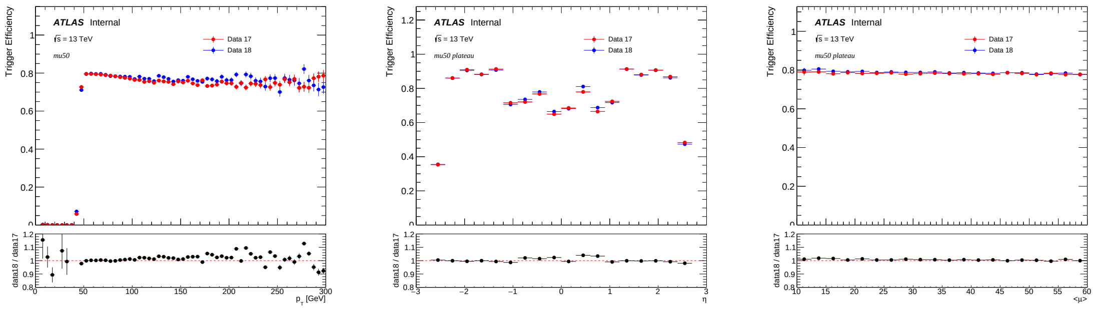
In Figure A.7 the efficiency curves as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_e60_lhmedium_nod0 are shown.

In Figure A.8 the efficiency curves as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_e140_lhloose_nod0 are shown. The trigger efficiency evaluated with 2017 and 2018 data show a very good agreement for all the trigger thresholds considered.



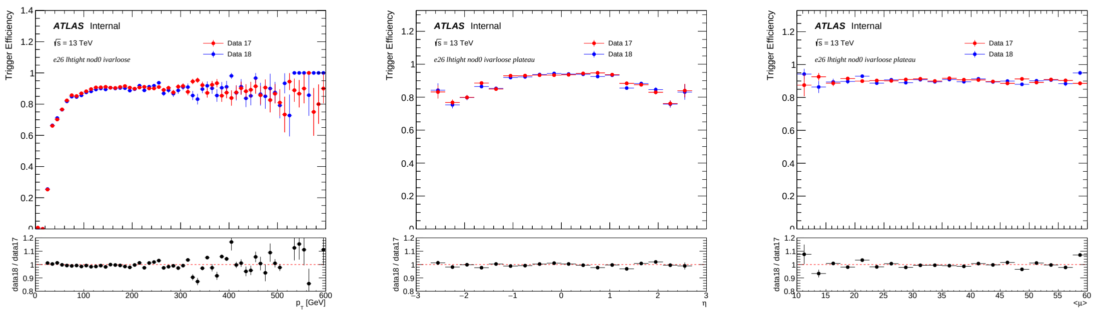
(a) (b) (c)

Figure A.4: Efficiency as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_mu26_ivarmedium. The efficiency in η and $< \mu >$ are taken at the trigger threshold at plateau.



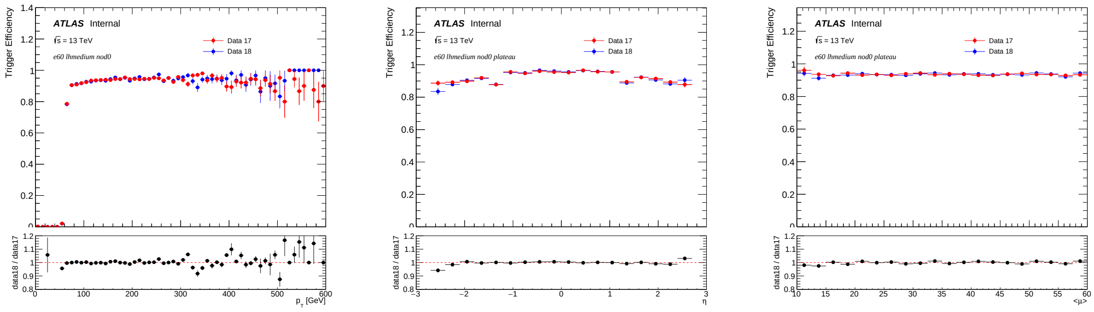
(a) (b) (c)

Figure A.5: Efficiency as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_mu50. The efficiency in η and $< \mu >$ are taken at the trigger threshold at plateau.



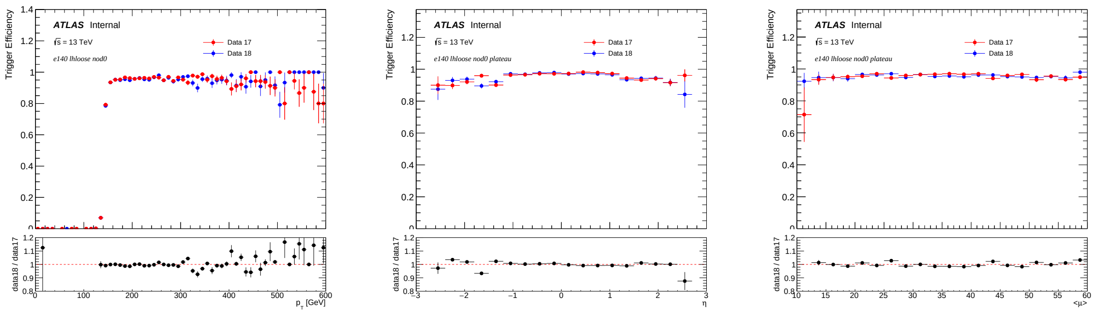
(a) (b) (c)

Figure A.6: Efficiency as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($< \mu >$) (c) for the trigger threshold HLT_e26_lhtight_nod0_ivarloose. The efficiency in η and $< \mu >$ are taken at the trigger threshold at plateau.



(a) (b) (c)

Figure A.7: Efficiency as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($<\mu>$) (c) for the trigger threshold HLT_e60_lhmedium_nod0. The efficiency in η and $<\mu>$ are taken at the trigger threshold at plateau.



(a) (b) (c)

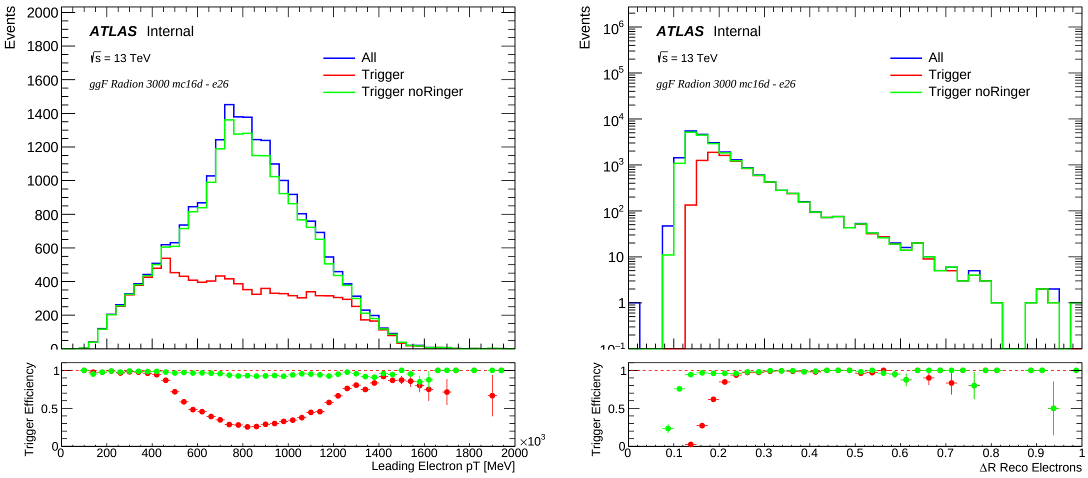
Figure A.8: Efficiency as a function of the p_T (a), η (b) and of the average number of interactions per bunch crossing ($<\mu>$) (c) for the trigger threshold HLT_e140_lhloose_nod0. The efficiency in η and $<\mu>$ are taken at the trigger threshold at plateau.

Appendix B

Appendix: Trigger at high-p_T based on ML in ATLAS

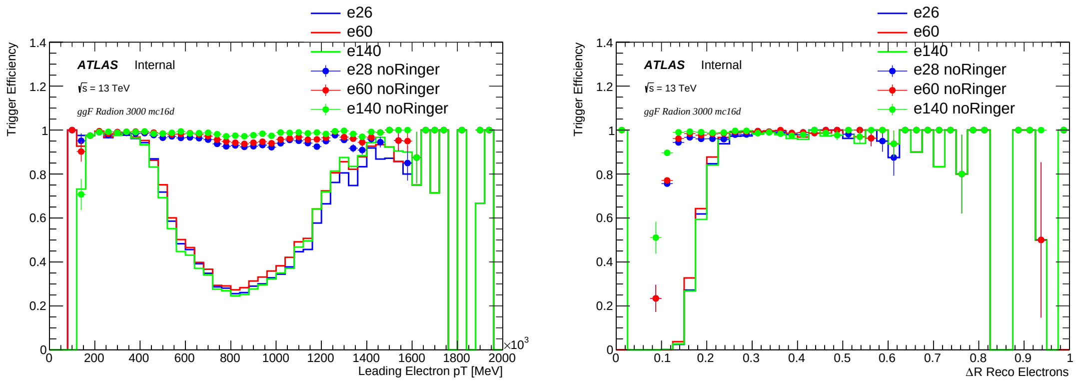
During Run-II new trigger item have been introduced; in particular, single-electron trigger based on Machine Learning approach, based on the so called *Ringer algorithm* [122], has been introduced in the 2017 data taking period. The Ringer algorithm increases the time taken by the fast calorimeter reconstruction step to 1–2 ms per event, approximately 45% slower than the cut-based algorithm. However, it reduces the number of input candidates for the more CPU-demanding fast tracking step (which takes about 64 ms per event) by a factor of 1.5–6. This factor depends on the detailed trigger configuration. Overall, the use of the Ringer algorithm enabled at least a 50% reduction in the CPU demand for the lowest-threshold unprescaled single-electron trigger.

Figures B.1 show the distributions of the leading electron p_T and of the ΔR of the electrons pair before and after the trigger requirement, as an example the lowest single electron item is showed. Figures B.2 show the trigger curves for all the single electron items as a function of the leading electron p_T and of the ΔR of the electrons pair. The distributions have perfomed using the simulated MC sample of a ggF spin-0 radion production with mass of 3000 GeV; the effect of the ringer algorithms have not never been observed in the data because it affect only the high-p_T region (> 500 GeV). The ringer algorithms, indeed, show trigger inefficiency at high-p_T region when the two electrons are close to each other.



(a) (b)

Figure B.1: Distribution of the leading electron pT (a) and $DeltaR$ between the two electrons (b) before and after the trigger requirement for the lowest single electron trigger item; in particular, ringer and no-ringer algorithms are compared.



(a) (b)

Figure B.2: Trigger curves for all the single electron item used as a function of the electron p_T (a) and ΔR of the electrons pair (b); ringer (lines) and no-ringer (dots) are compared.

Appendix C

Appendix: Signal Efficiency

The acceptance times efficiency for the signal region selection is defined as:

$$A \cdot \epsilon = \frac{N^{events}_{selected}}{N^{events}_{Collected\ by\ ATLAS}} \cdot \frac{N^{events}_{Collected\ by\ ATLAS}}{N^{events}_{produced}} = \frac{N^{events}_{selected}}{N^{events}_{produced}} \tag{C.1}$$

The distributions of $A \cdot \epsilon$ as a function of the resonance mass for the different spin models are shown in Figures C.1-C.2-C.3-C.4 for the 0-lepton channel, C.5-C.6-C.7-C.8 for the 1-lepton channel, C.9-C.10-C.11-C.12 for the 2-lepton channel,.

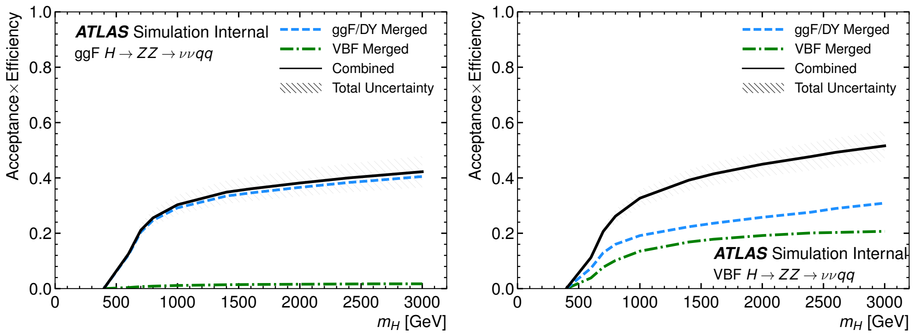


Figure C.1: Selection acceptance times efficiency for the $H \to ZZ \to \nu\nu qq$ events from MC simulations as a function of the Higgs boson mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \to \nu\nu J$ selection.

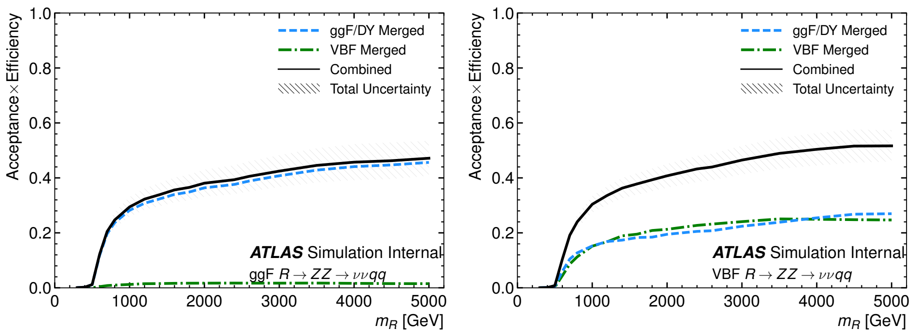


Figure C.2: Selection acceptance times efficiency for the $R \to ZZ \to \nu\nu qq$ events from MC simulations as a function of radion mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \to \nu\nu J$ selection.

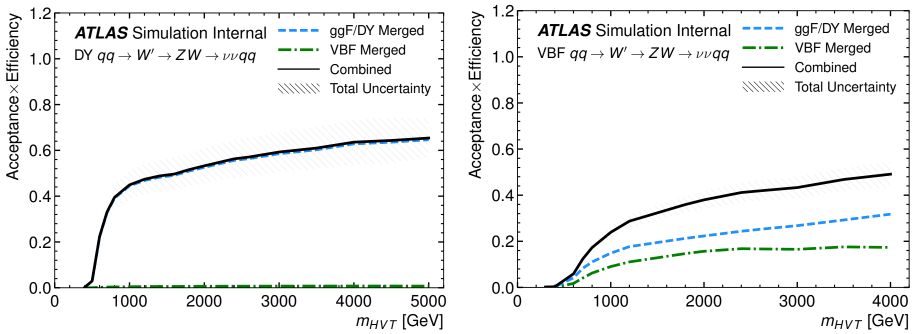


Figure C.3: Selection acceptance times efficiency for the $W' \rightarrow ZW \rightarrow \nu\nu qq$ events from MC simulations as a function of W' mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \rightarrow \nu\nu J$ selection.

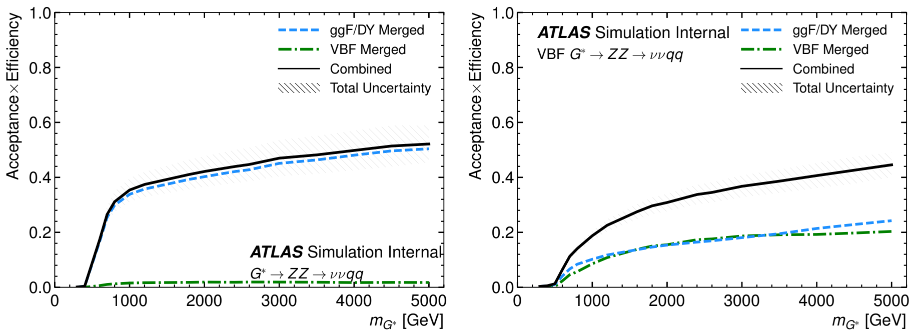


Figure C.4: Selection acceptance times efficiency for the $G \rightarrow ZZ \rightarrow \nu\nu qq$ events from MC simulations as a function of radion mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \rightarrow \nu\nu J$ selection.

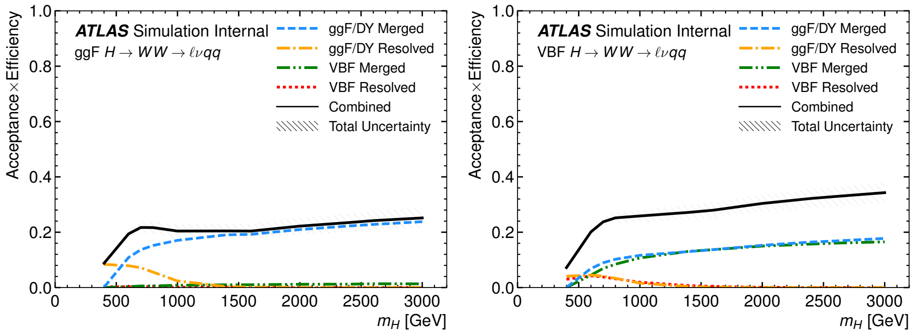


Figure C.5: Selection acceptance times efficiency for the $H \to WW \to \ell\nu qq$ events from MC simulations as a function of the Higgs boson mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $WV \to \ell\nu J$ selection and the resolved regions of the $WV \to \ell\nu jj$ selection.

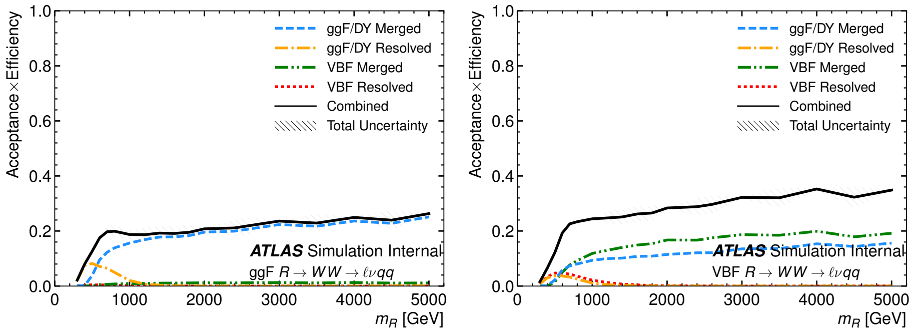


Figure C.6: Selection acceptance times efficiency for the $R \to WW \to \ell\nu qq$ events from MC simulations as a function of the radion mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $WV \to \ell\nu J$ selection and the resolved regions of the $WV \to \ell\nu jj$ selection.

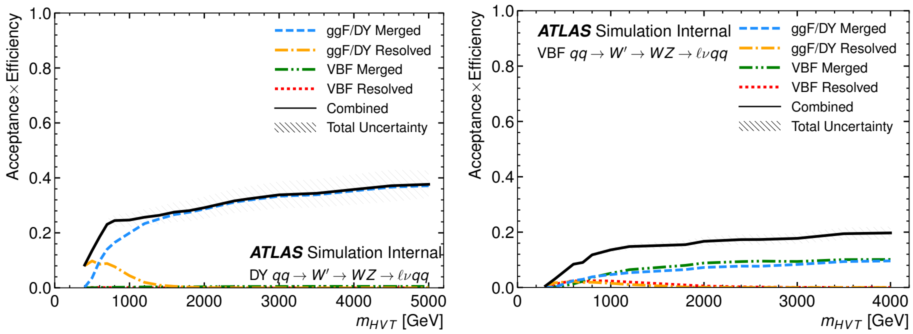


Figure C.7: Selection acceptance times efficiency for the $W' \to WZ \to \ell\nu qq$ events from MC simulations as a function of the W' mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $WV \to \ell\nu J$ selection and the resolved regions of the $WV \to \ell\nu jj$ selection.

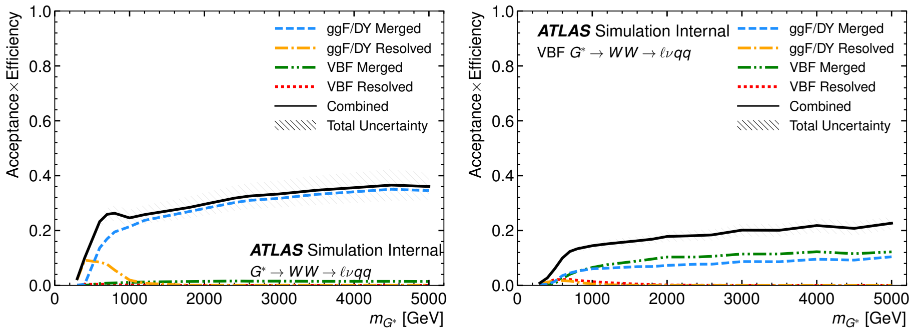


Figure C.8: Selection acceptance times efficiency for the $G \to WW \to \ell\nu qq$ events from MC simulations as a function of the Graviton mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $WV \to \ell\nu J$ selection and the resolved regions of the $WV \to \ell\nu jj$ selection.

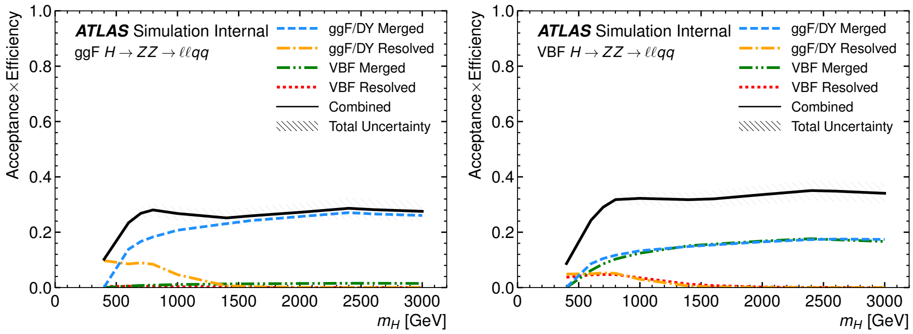


Figure C.9: Selection acceptance times efficiency for the $H \to ZZ \to \ell\ell qq$ events from MC simulations as a function of the Higgs boson mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \to \ell\ell J$ selection and the resolved regions of the $ZV \to \ell\ell jj$ selection.

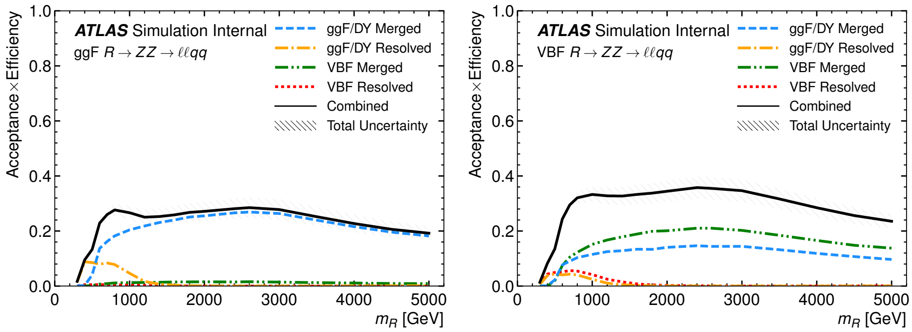


Figure C.10: Selection acceptance times efficiency for the $R \to ZZ \to \ell\ell qq$ events from MC simulations as a function of the radion mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \to \ell\ell J$ selection and the resolved regions of the $ZV \to \ell\ell jj$ selection.

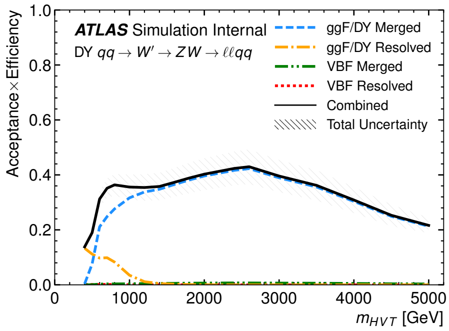


Figure C.11: Selection acceptance times efficiency for the $W' \to ZW \to \ell\ell qq$ events from MC simulations as a function of the W' mass for (a) Drell-Yan and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \to \ell\ell J$ selection and the resolved regions of the $ZV \to \ell\ell jj$ selection.

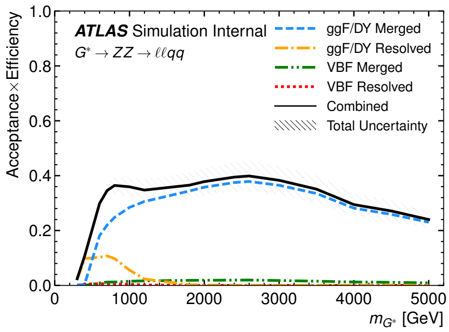


Figure C.12: Selection acceptance times efficiency for the $G \to ZZ \to \ell\ell qq$ events from MC simulations as a function of the Graviton mass for (a) ggF and (b) VBF production, combining the merged HP and LP signal regions of the $ZV \to \ell\ell J$ selection and the resolved regions of the $ZV \to \ell\ell jj$ selection.

Appendix D

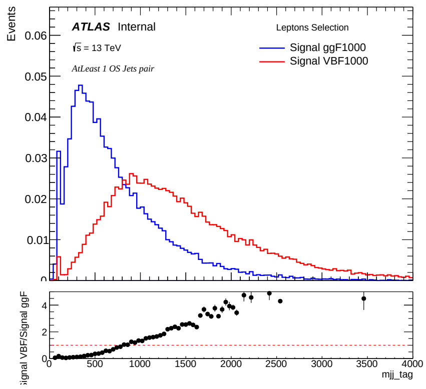
Appendix: Low-level VS high-level input variables for NN approach

Discussion about input features: low VS high level variables

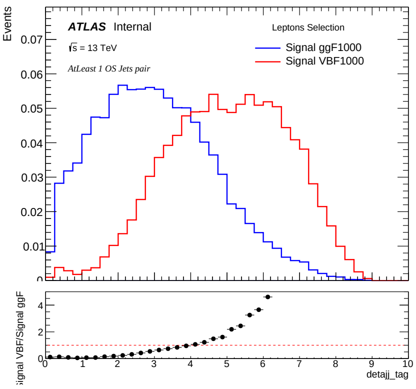
The choice of the variables to use is a crucial point during the development of a new approach. As general result of the deep learning a set of features that represent basic information of the problem (such as information coming directly form the detectors) implies better performances respect a set of features built combining basic information if a deep learning approach is used. In high-energy physics, we can define:

- *Low-level variables*: simple and basic variables, such as the components of the 4-momentum of the reconstructed objects;
- *High-level variables*: each variable should be constructed starting from the 4-momentum.

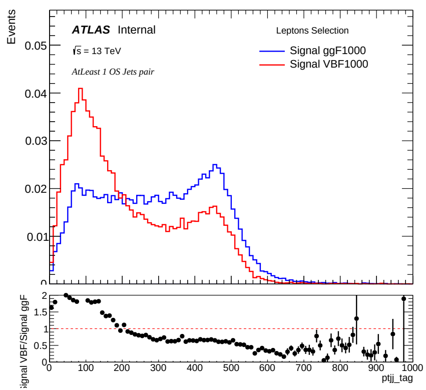
Some structural variables related to the 2 tag jets have been built, such as the invariant mass of the jets pair, m_{jj}^{Tag}, the pseudo-rapidity separation of the 2 jets, $|\Delta\eta_{jj}^{Tag}|$, the transverse momentum of the jets pair, $p_{T\ jj}^{Tag}$ and the higher η of the 2 jets; the shapes distributions for the VBF and the ggF signals are showed in Figures D.1.



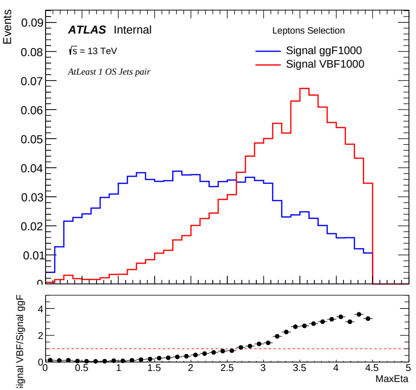
(a) m_{jj}^{Tag}



(b) $|\Delta\eta_{jj}^{Tag}|$



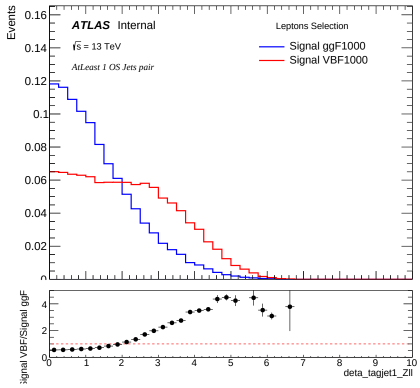
(c) p_{Tjj}^{Tag}



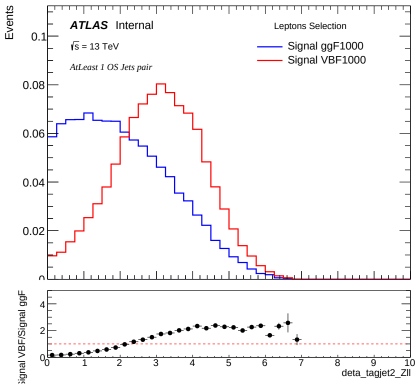
(d) $max(\eta_{TagJet1}, \eta_{TagJet2})$

Figure D.1: Comparison of shapes distribution for the structural variables considered.

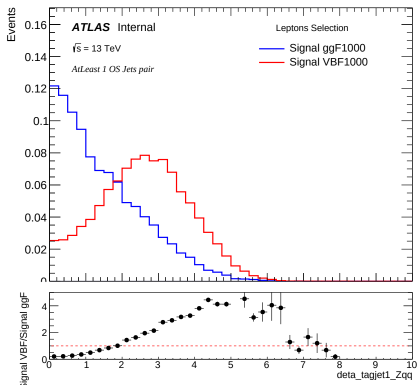
Structural variables considering the η separation between the 2 Z bosons coming from the resonance and the 2 VBF jets have been also built. For the hadronic Z a very simple selection has been applied in order to not reduce the events for the training of the net. These variables are showed in Figures D.2.



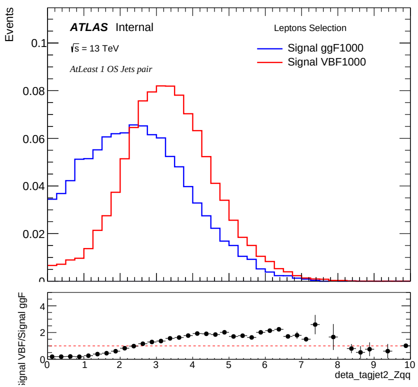
(a) $|\Delta\eta(Z_{ll}, jet_1^{Tag})|$



(b) $|\Delta\eta(Z_{ll}, jet_2^{Tag})|$



(c) $|\Delta\eta(Z_{qq}, jet_1^{Tag})|$

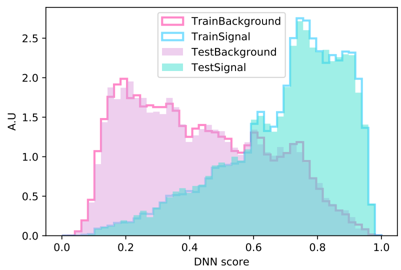


(d) $|\Delta\eta(Z_{qq}, jet_2^{Tag})|$

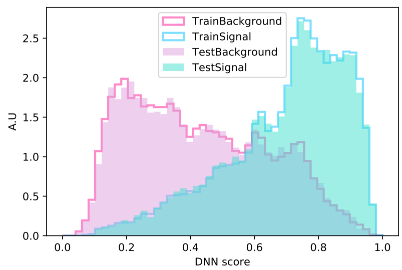
Figure D.2: Comparison of shapes distribution for the structural variables considered.

The DNN training has been performed using the structural variables; in particular, two cases have been considered, the first considering 6 variables, m_{jj}^{Tag}, $|\Delta\eta_{jj}^{Tag}|$, $p_{T\ jj}^{Tag}$, $max(\eta_{TagJet1}, \eta_{TagJet2})$, $|\Delta\eta(Z_{ll}, jet_1^{Tag})|$, $|\Delta\eta(Z_{ll}, jet_2^{Tag})|$, and a second with 8 variables adding also $|\Delta\eta(Z_{qq}, jet_1^{Tag})|$ and $|\Delta\eta(Z_{qq}, jet_2^{Tag})|$.

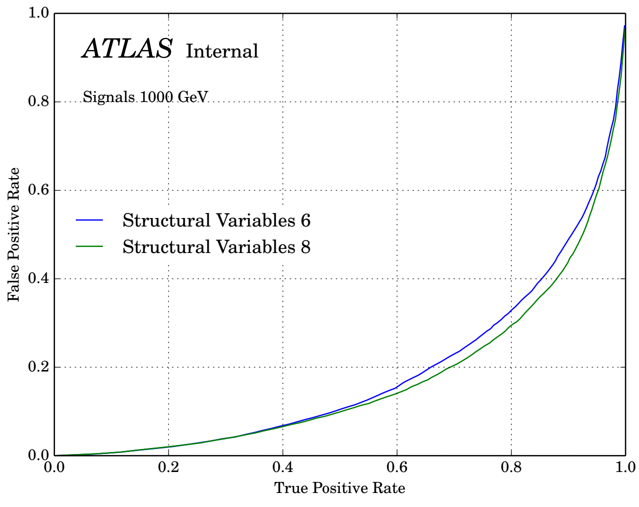
The training and test scores are showed in Figure D.3 (a)-(b), while the ROC curves comparison is in Figure D.3 (c).



(a) *Set of 6 structural variables.*



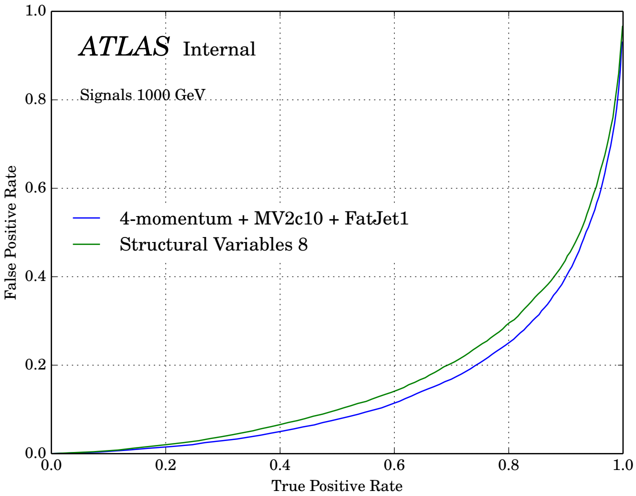
(b) *Set of 8 structural variables.*



(c)

Figure D.3: DNN performances using the Structural Variables.

A comparison of the DNN performances using the low-level variables and the high-level variables is showed in Figure D.4. The performances of the net are better for the low-level variables; the gain we have using the low-level variables is reported on table D.1 for some WP. With this kind of network architecture, it seems that use more simple information give the chance to improve the VBF/ggF separation respect than more sophisticated information.



(a)

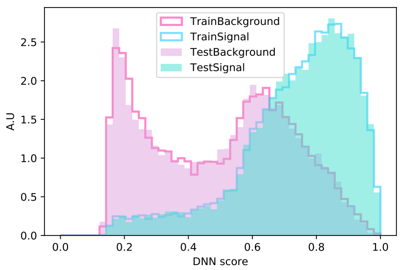
Figure D.4: Comparison of the DNN performances using the 4-momentum VS the Structural Variables approaches.

	False Positive Rate		
True Positive Rate	8 Structural Variables	6jets 4-momentum + MV2c10 + FatJet1 4-momentum	False Positive decrease
13%	59%	63%	7%
40%	6.7%	4.8%	40%
60%	14%	12%	17%
80%	30%	25%	20%

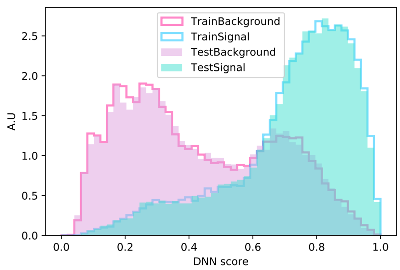
Table D.1: False positive rate for different True Positive Rate WP with High-level and low-level variables; the last column represent the gain in terms of the False Positive decrease.

Performances varying the signal mass

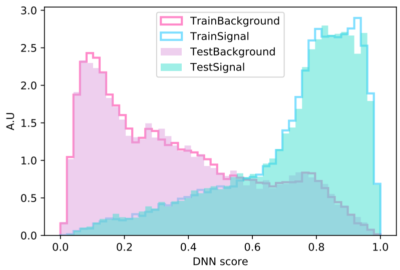
The DNN training has been performed also considering all the mass point available, $300, 700, 1000, 2000, 3000\ GeV$. The scores are showed in Figure D.5 for the case with the 4-momentum variables while in Figure D.6 there are the scores in the case with the structural variables.



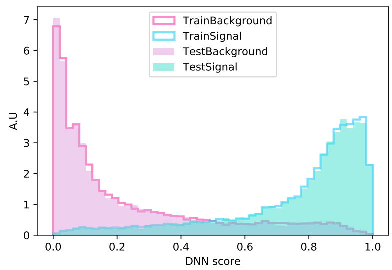
(a)



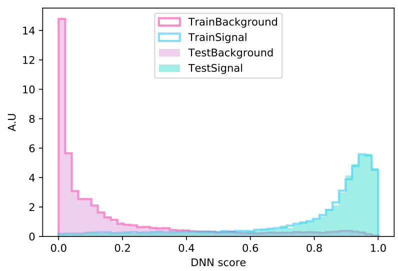
(b)



(c)

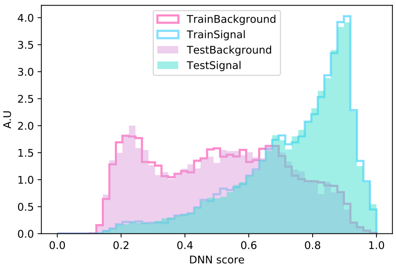


(d)

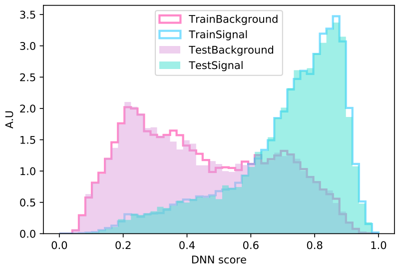


(e)

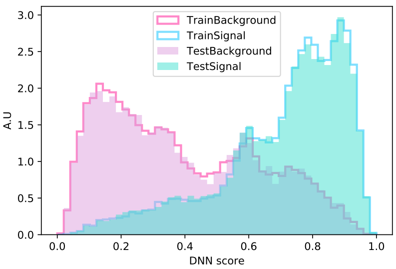
Figure D.5: DNN scores with the 4-momentum approach varying the signal mass value.



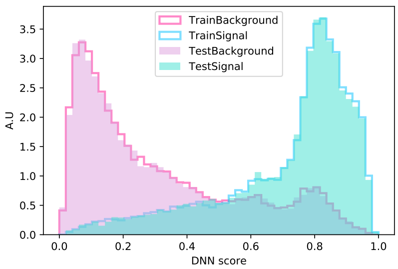
(a)



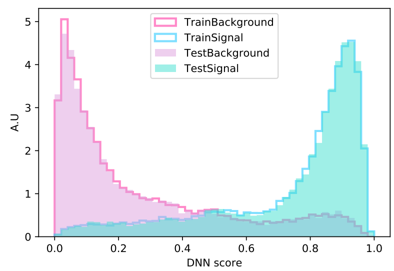
(b)



(c)



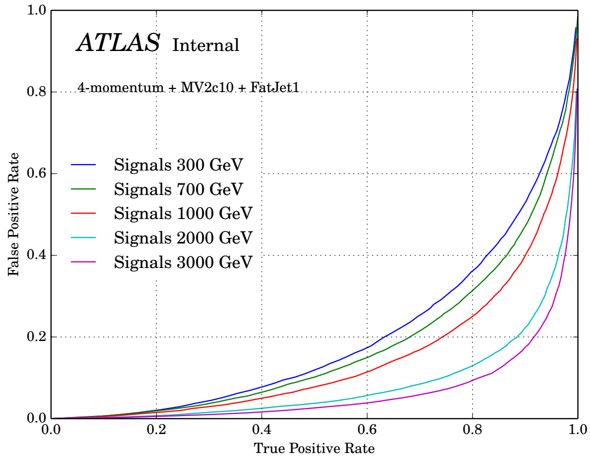
(d)



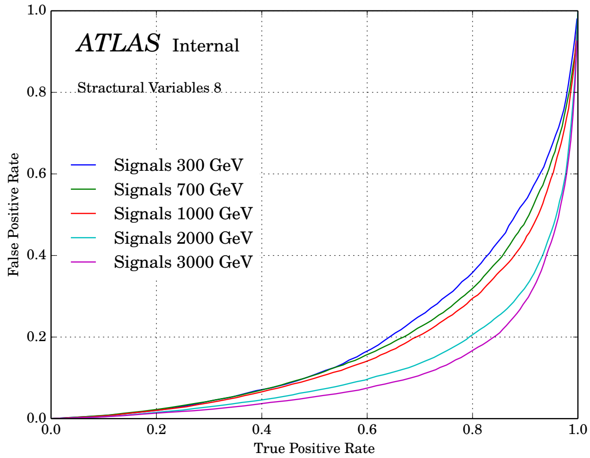
(e)

Figure D.6: DNN scores with the 4-momentum approach varying the signal mass value.

In Figure D.7 the ROC curve obtained varying the signal mass are showed. The VBF/ggF separation power of the DNN increase with the value of the resonance mass and the performances for the low-level variables are better than for the high-level variables for each mass point.



(a)



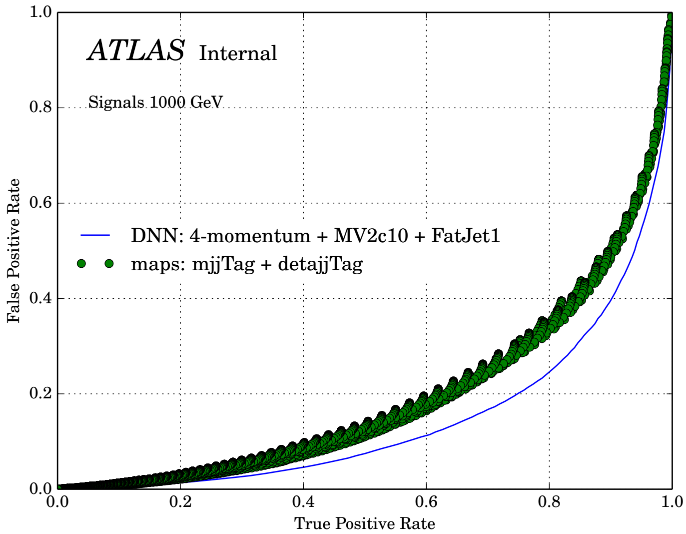
(b)

Figure D.7: DNN performances varying the signal mass value in the range $[300, 3000]$ GeV using the 4-momentum (a) and the Structural variables approach (b).

Classification performance comparison with the 2-dimensional approach

The efficiency selection varying the values of the cuts on the m_{jj}^{Tag} and $|\Delta\eta_{jj}^{Tag}|$ variables has been evaluated for both VBF and ggF 1000 GeV samples. Each pair of values of the VBF efficiency and ggF efficiency corresponds to a point in the
$True\ Positive\ VS\ False\ Positive$ plane. The comparison of the all possible WP for the 2-dimensional cut based analysis and the DNN approach is showed in Figure D.8(c).

The gain we obtained is up to $50 - 60\%$ in terms of the decrease of the False Positive Rate and up to $\sim 30\%$ in terms of the increasing of the True Positive Rate (Table F.4).



(a)

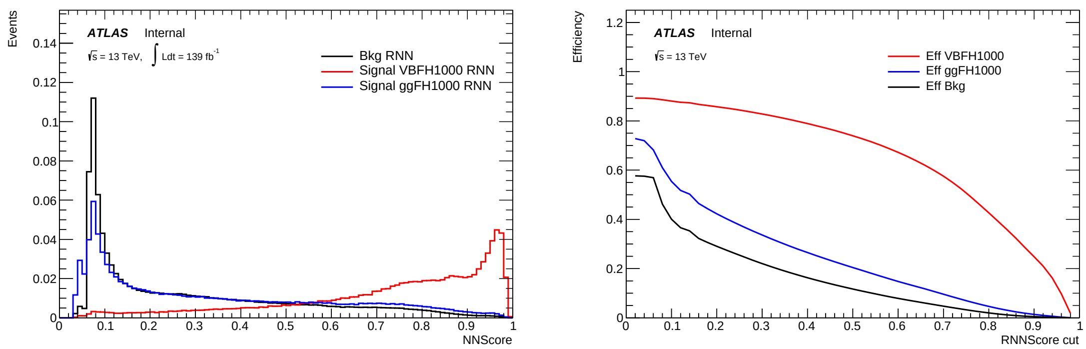
Figure D.8: ROC curve comparison for the 2-dimensional approach and for the new 4-momentum approach with DNN.

Appendix E

Appendix: RNN setting for the analysis

E.1 Choice of the WP for the categorisation

Figure E.1a shows the RNN score distributions for the VBF and ggF 1 TeV signals and the SM backgrounds (including Z+jets, diboson, top and W+jets as usual). We can find good separation of the VBF signal from not only the ggF signal but also the SM background. Figure E.1b shows the efficiency to signal and background as a function of threshold on RNN score. In this section, the optimization study of the cut on the RNN score is discussed, to maximize the VBF signal sensitivity against the SM background.



(a) (b)

Figure E.1: Shapes distribution of the RNN scores (a) and efficiency curves for VBF and ggF signals and for the backgrounds.

In order to find the optimal RNN score cut several estimator have been considered.

Estimator A $\epsilon_{\mathrm{VBF}} \cdot (1 - \epsilon_{\mathrm{ggF}})$,

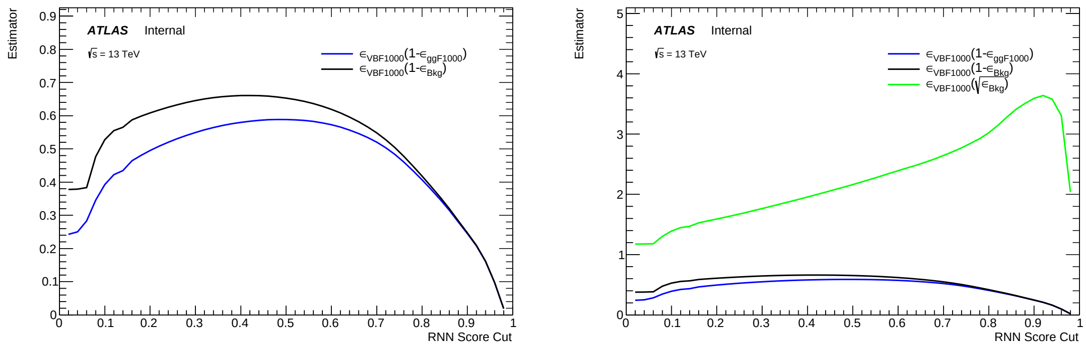
Estimator B $\epsilon_{\mathrm{VBF}} \cdot (1 - \epsilon_{\mathrm{bkg}})$,

Estimator C $\epsilon_{\mathrm{VBF}} / \sqrt{\epsilon_{\mathrm{bkg}}}$,

where ϵ_S with $S = VBF, ggF, bkg$ is defined as follow:

$$\epsilon_S = \frac{\#\ of\ events\ of\ the\ process\ S\ with\ RNN > X}{\#of\ events\ of\ the\ process\ S} \tag{E.1}$$

The first estimator is maximized if the efficiency to VBF signal is moderately high and the ggF signal contamination is relatively suppressed. The second estimator is defined by replacing ϵ_{ggF} of the first to ϵ_{bkg}, to find the best compromise for the signal efficiency and the background rejection. The third estimator is used to evaluate the relative sensitivity gain using the cut on RNN score.



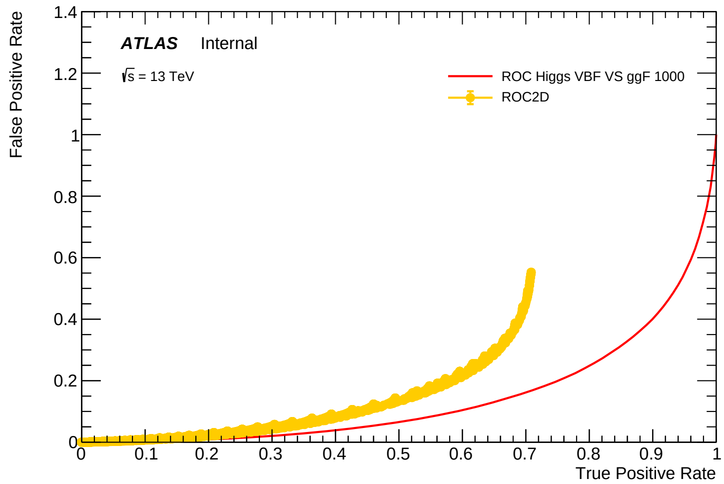
(a) (b)

Figure E.2: Several estimator considered varying the RNN Score cut.

Three different working points (WPs) have been tested, i.e. $RNN > 0.6$, 0.7 and 0.8. Figure E.4 shows the expected upper limits to VBF signals from merged HP, LP and resolved categories, separately, for three working points. Flat 10% systematic uncertainty on the background normalization is considered for these plots. With all three working points, the sensitivity is significantly increased from the cut-based approach. A tighter cut on RNN score is favored at the lower mass region, since the higher background rejection is important, while a looser cut is preferred at the high mass region. Due to the extremely small cross sections of VBF signals, the lower mass region is still important in this analysis. Therefore, we can give a priority to the sensitivity in low-mass region.

We also need to think about the ggF signal contamination in the VBF category. With looser WPs, not only the efficiency to VBF signals but also the ggF signal contamination is increased. Consequently, the sensitivity to ggF signal in the ggF category is decreased. Figure E.5 shows the expected upper limits to ggF signals with the different three WPs. A tighter cut is clearly preferred in terms of the ggF signal contamination.

At the end, the working point is chosen to give the same background rejection in VBF category as the cut-based analysis. In 2-lepton channel, the cut to achieve that is $RNN > 0.78$, as shown in Figure E.6. For the simplicity, $RNN > 0.8$ is used to select the VBF event candidates. Table E.1 shows the efficiency of $RNN > 0.8$ for each mass point, compared to the cut-based analysis and Figure E.3 shows the ROC curves for the RNN and cut based approach for the 1000 GeV signals. Using the cut-based VBF categorization, efficiency to VBF signals strongly depends on the signal mass (37% at 700 GeV and 27% at 3 TeV). Using the RNN instead, we can suppress the mass dependency and find almost flat 38–44% efficiency at each mass point. Figure shows the relative improvement of the VBF signal efficiency and the ggF signal contamination in the new VBF category compared to the cut-based. We can get 0-60% more VBF signals in the VBF category depending on mass; while ggF signal contamination is suppressed by about 20%.

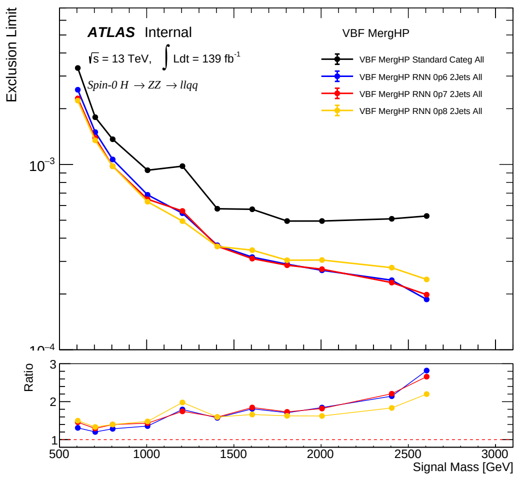


(a)

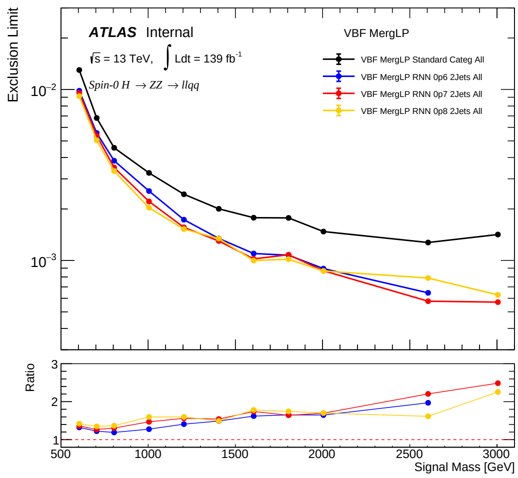
Figure E.3: ROC curves for the RNN approach (red line) and cut based approach (yellow dots).

Process	Yields at Dileptons Selection	Efficiency RNN > 0.6	Efficiency RNN > 0.7	Efficiency RNN > 0.8	Cut-based VBF category	RNN 0.8 / Cut-based
Backgrounds	10364900	7.9%	4.7%	2.0%	2.5%	0.8
VBF H 700 GeV	504	62.2%	52.3%	38.2%	37.3%	1.02
VBF H 1000 GeV	229	67.2%	57.6%	42.7%	37.8%	1.13
VBF H 3000 GeV	2.1	69.6%	60.9%	44.2%	27.0%	1.64
ggF H 700 GeV	1373	14.2%	9.0%	4.2%	5.5%	0.76
ggF H 1000 GeV	486	14.8%	9.6%	4.6%	6.1%	0.76
ggF H 3000 GeV	2.4	15.7%	10.3%	5.3%	6.3%	0.84

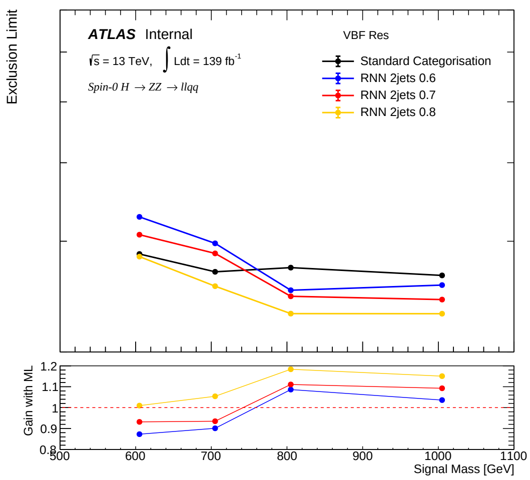
Table E.1: Yields and efficiency for the different RNN WPs.



(a)

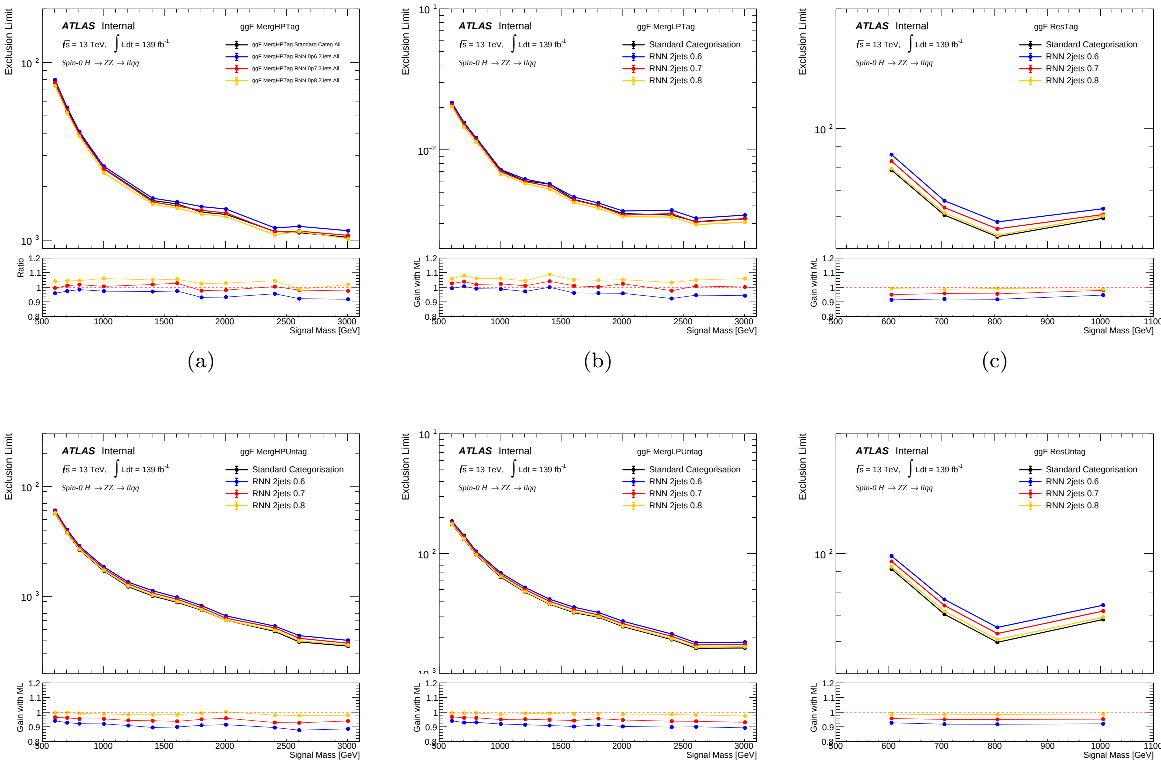


(b)



(c)

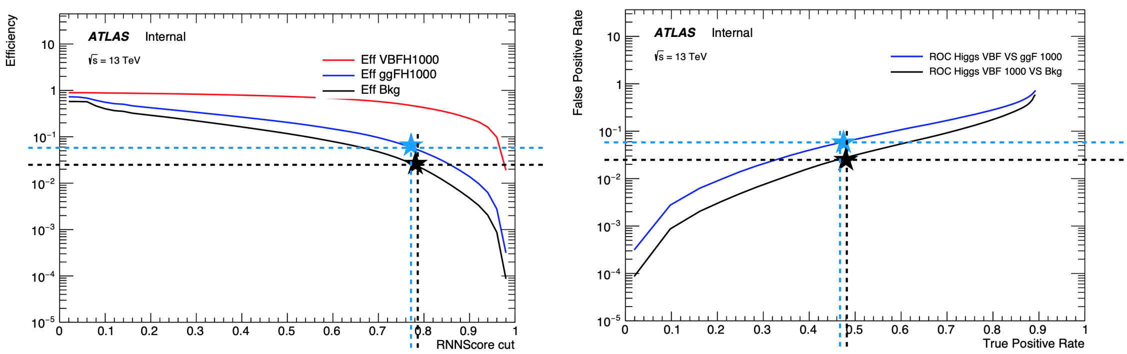
Figure E.4: Expected exclusion limits to VBF signals from the VBF SRs. Results from three WPs for RNN cut (colored) and the cut-based analysis (black) are shown. Flat 10% systematic uncertainty on the background normalization is assumed. Bottom panel shows the ratio of the cut-based sensitivity to the sensitivities of RNN analysis using different WPs (> 1 is better).



(a) (b) (c)

(d) (e) (f)

Figure E.5: Expected exclusion limits to ggF signals from the ggF SRs; (a) merged HP tagged (b) merged LP tagged (c) resolved tagged (d) merged HP untagged (e) merged LP untagged and (f) resolved untagged categories. Results from three WPs for RNN cut (colored) and the cut-based analysis (black) are shown. Flat 10% systematic uncertainty on the background normalization is assumed. Bottom panel shows the ratio of the cut-based sensitivity to the sensitivities of RNN analysis using different WPs (> 1 is better).



(a) (b)

Figure E.6: (a) Efficiencies to VBF and ggF signals and the background as a function of threshold of the RNN score. (b) ROC curves to compare the efficiency to VBF signal with ggF signal (blue) and to compare VBF signal with the SM background (black). The stars show the performances of the cut-based analysis.

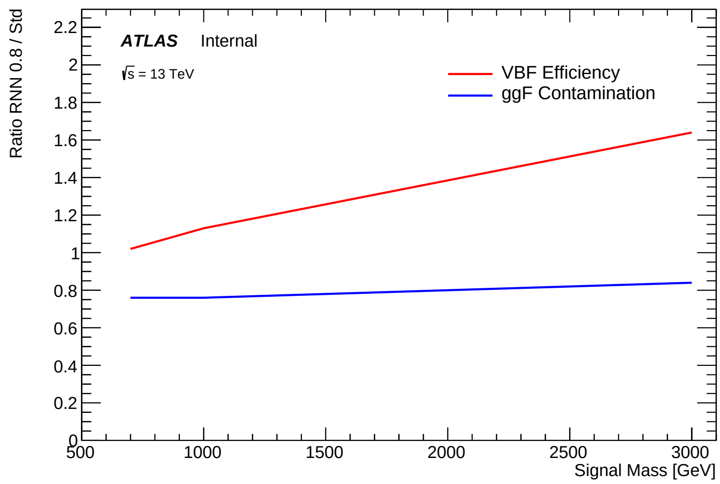
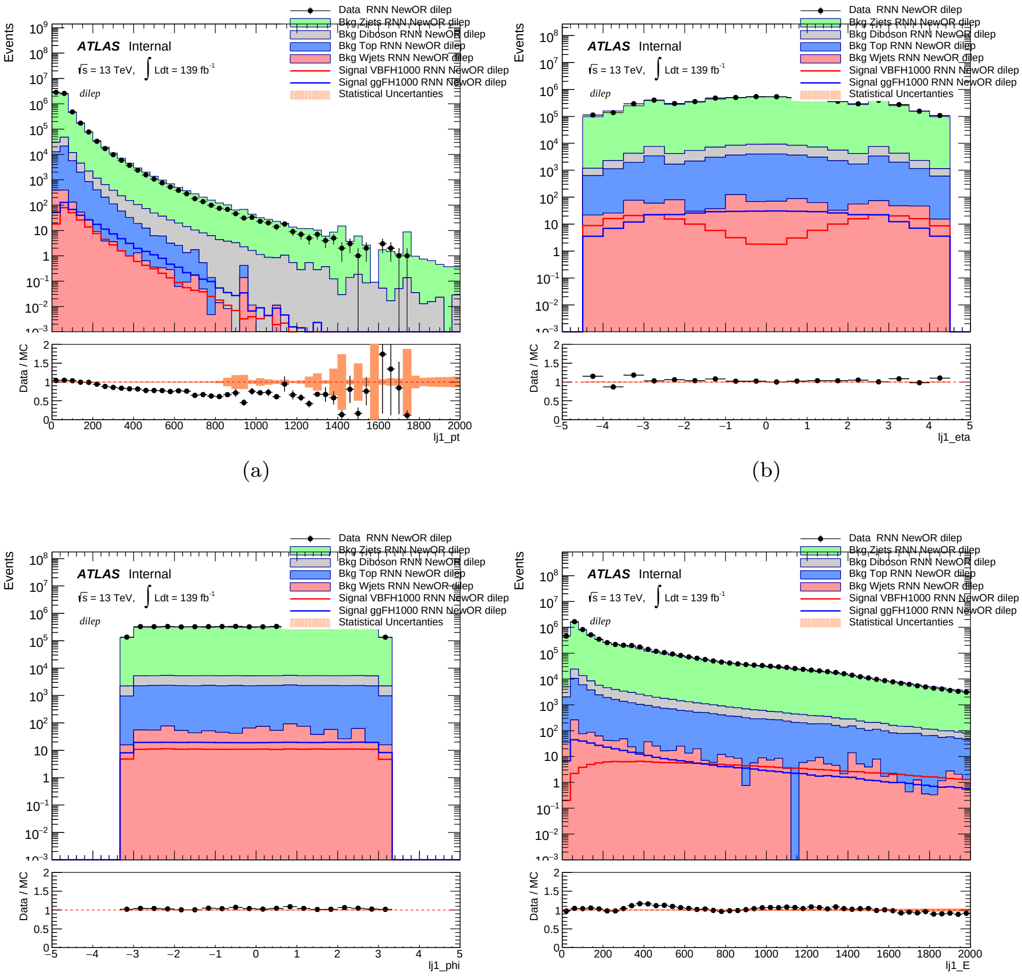


Figure E.7: Ratio of number of VBF (red) and ggF signals (blue) in VBF category defined by $RNN > 0.8$ to that in the cut-based VBF category.

E.2 Modeling of input variables

Figures E.8-E.9 shows the data/MC agreements for the leading and subleading jet 4-momenta used as RNN inputs, after the preselection. The data shows a good agreement with MC.

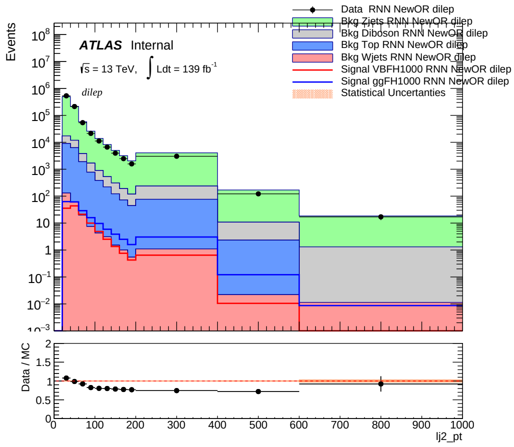


(a) (b) (c) (d)

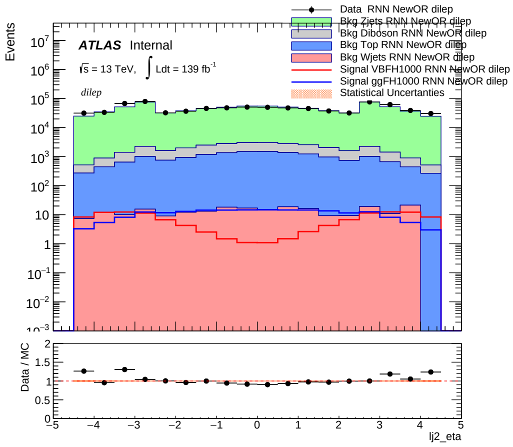
Figure E.8: Data/MC plots of 4-momentum of the Jet1 used as RNN input.

Figure E.10 shows the RNN score data/MC comparison. The data/MC agreement is good on the full range of the score values.

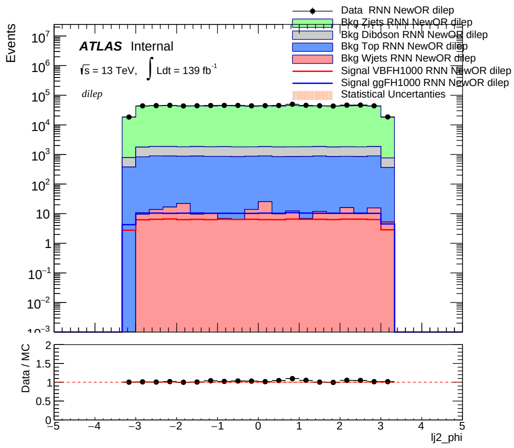
Figures E.11-E.12 shows the data/MC agreements for the leading and subleading jet 4-momenta used as RNN inputs, after the preselection; this time all the systematic uncertainties described in Section 5.10 are considered in the ratio plot. The data shows a good agreement with MC within the systematic variation bands. Figure E.13 shows the RNN score data/MC comparison. The data/MC agreement is good on the full range of the score values and within the systematic uncertainties.



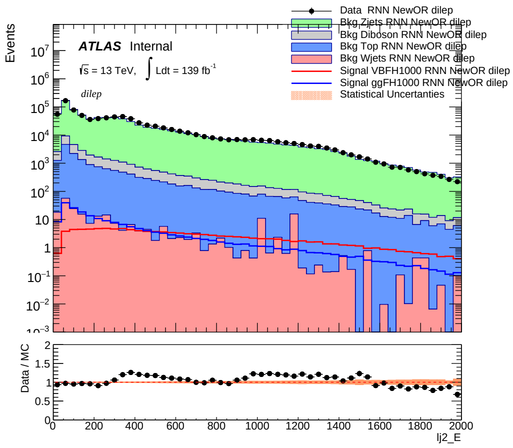
(a)



(b)

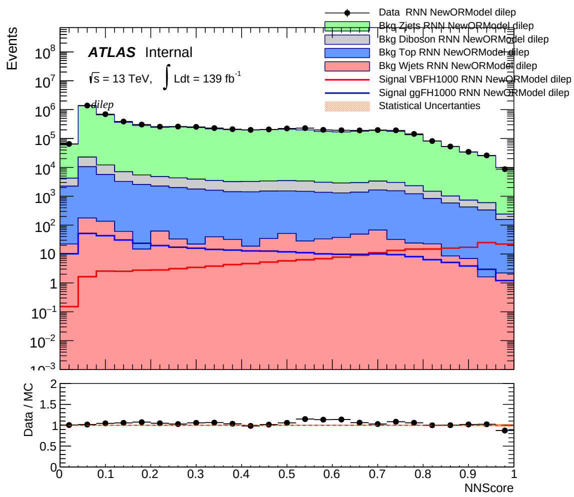


(c)



(d)

Figure E.9: Data/MC plots of 4-momentum of the Jet2 used as RNN input.



(a)

Figure E.10: Data/MC plots of RNN Score.

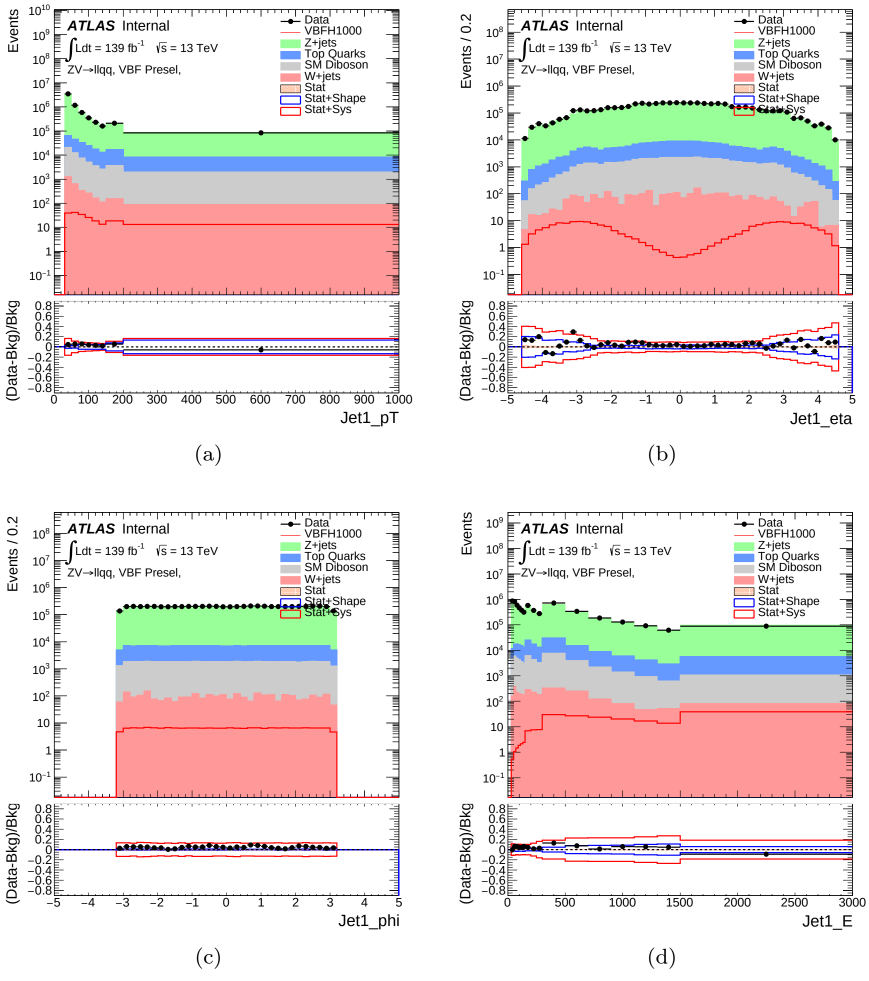
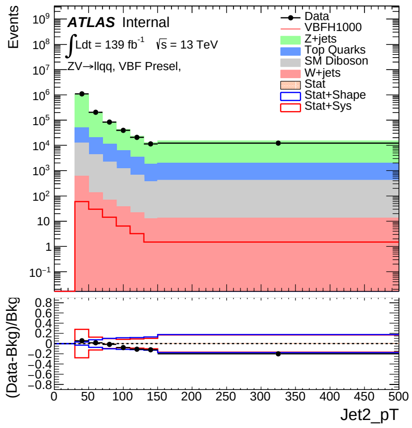
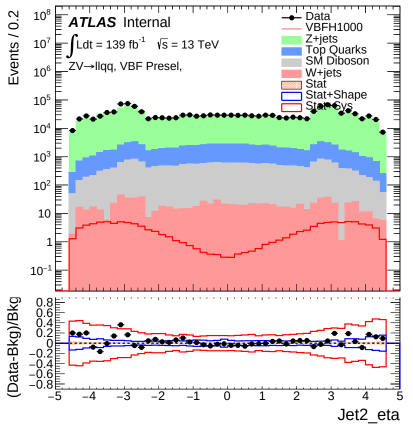


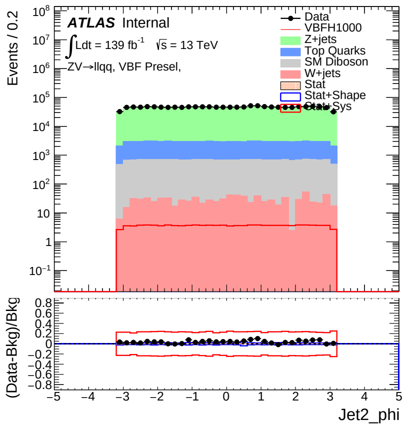
Figure E.11: Data/MC plots of 4-momentum of the Jet1 used as RNN input; systematic are taken into account in the ratio plot.



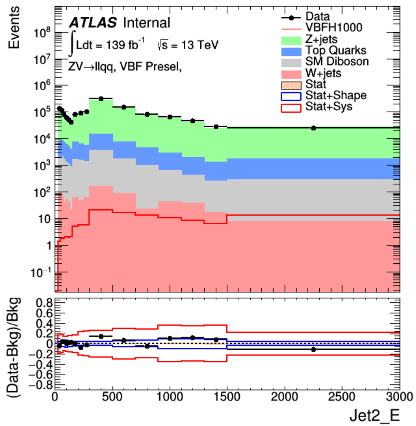
(a)



(b)

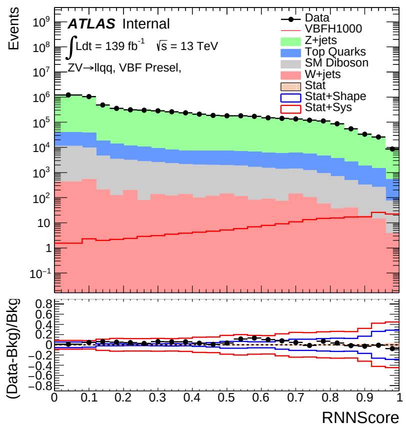


(c)



(d)

Figure E.12: Data/MC plots of 4-momentum of the Jet2 used as RNN input; systematic are taken into account in the ratio plot.



(a)

Figure E.13: Data/MC plots of the RNN Score; systematic are taken into account in the ratio plot.

E.3 Discussion about maximum number of jets

The first RNN model development exploited the possibility to have a maximum of 6 jets in the input sequence ($\sim 95\%$ of the signal events have less than 6 jets). The possibility to reduce the maximum number of jets has been investigated.

The figures E.14 show the scores shapes and the ROC curve performances varying the maximum number of input jets to the RNN. From 2 jets up to 6 jets the performances are very similar.

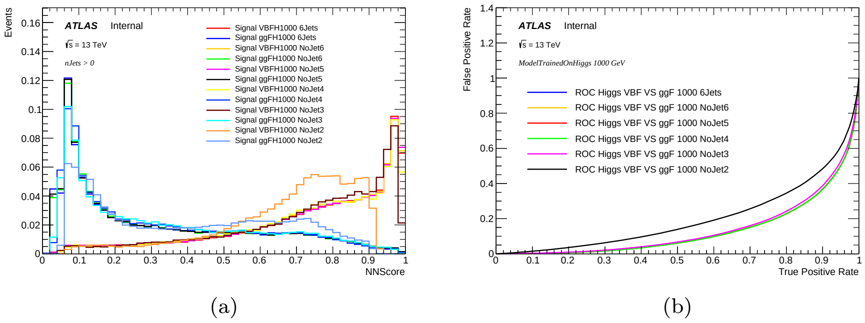


Figure E.14: RNN score shapes for VBF and ggF signals for different training performed with different maximum number of jets as input (a) and relative ROC curves (b).

The figure E.15 shows that the efficiencies for signals and backgrounds do not vary a lot if we move from the 6 jets to the 2 jets model.

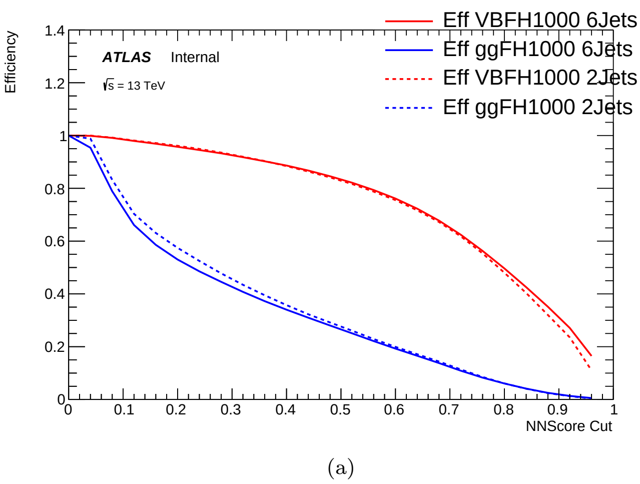
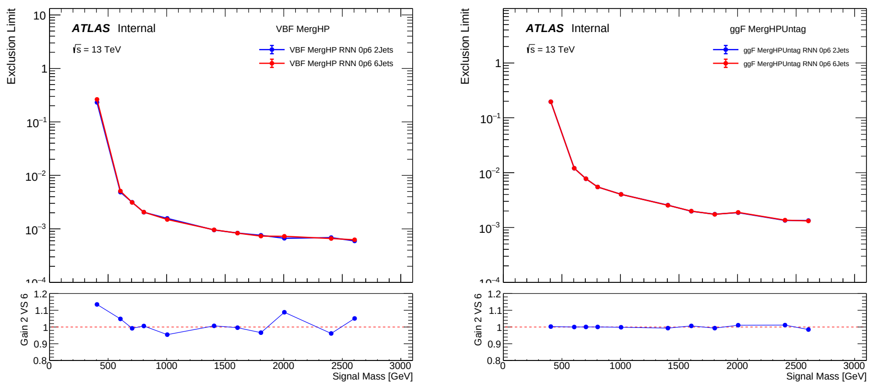


Figure E.15: Efficiency curves for VBF and ggF signals in the training setup with 6 jets as maximum (solid) and with 2 jets as maximum (dashed).

The figure E.16 shows the exclusion limits for 2 regions as an example, comparing the 2 RNN model, also the sensitivity are not affected if we reduced the input jets as a maximum of 2.



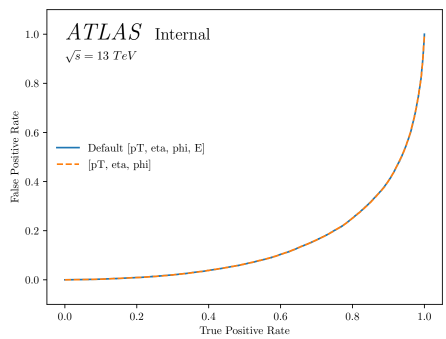
(a) (b)

Figure E.16: Expected upper limit on the production of the resonance comparing a scenario with the RNN trained with 6 jets as maximum (red) and 2 jets as maximum (blue).

After we observed that the performances do not decrease strongly when we move to only 2 jets, we decided to use this model in order to avoid any extra systematics coming from the modelling of an higher number of jets from the MC simulations.

E.4 Features Ranking with re-training

The RNN model has been re-trained removing the energy component and using the simple 3-momentum of the jets ($[p_T, \eta, \phi]$) in order to have a further check that the energy do not affect the RNN performances; the result is showed in Figure E.17. The 2 model are consistent.

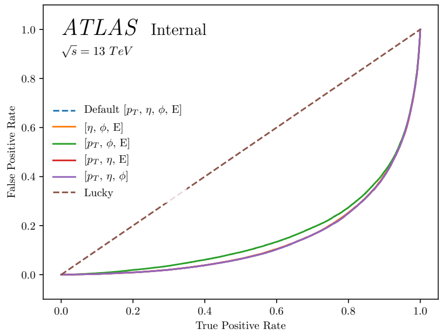


(a)

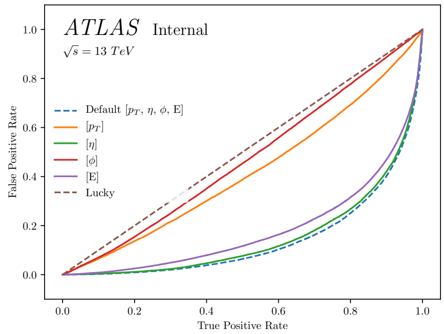
Figure E.17: ROC curves comparison for the 2 RNN model with and without the energy component.

In order to deeply investigate the energy role in the usage of the 4-momentum of the jets as RNN features for the VBF/ggF classification, the RNN model has been re-trained varying the input features. Firstly, the RNN model has been re-trained removing one features at the time, then the RNN model has been re-trained using just one input feature, the ROC curves comparison are reported in Figure E.18.

What we learn from these checks is:



(a)



(b)

Figure E.18: ROC curves comparison varying the input features set, using the N-1 variables (a) and using only 1 variable (b).

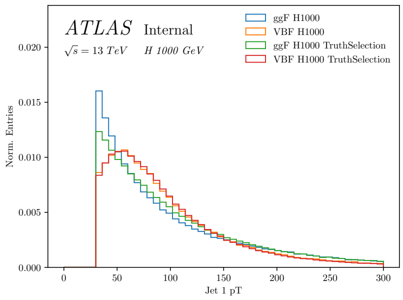
- the η component is the most powerful (right plot);
- p_T and E are equivalent if used with η, ϕ (left plot);
- removing η the net is able to recover it from the correlation between p_T and E.

This is true for this kind of problem but it could be different for other classification problems!

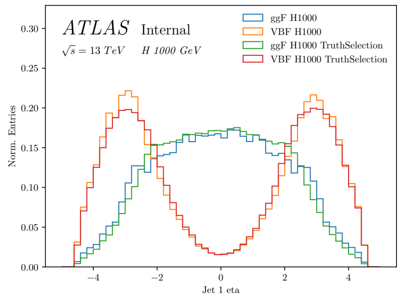
E.5 RNN at Truth Level

The RNN model has been tested at the truth level. Truth MC simulation has been considered without the ATLAS detector simulation. A loose pre-selection has been performed in order to select the RNN-input jets. Good leptons have been selected to come from the decay of Z/W bosons and not from the decay of hadrons or taus in the event and having $p_T > 30\ GeV$ and $|\eta| < 2.5$. Small-R jets have been selected to not overlap with good electrons ($\Delta R > 0.2$) and having $p_T > 30\ GeV$ and $|\eta| < 4.5$. Than, only events having a leptons pair in the Z boson mass window ($70 - 110\ GeV$) have been selected. The two RNN input jets have been selected according the selection procedure discussed in Section 5.6.1.

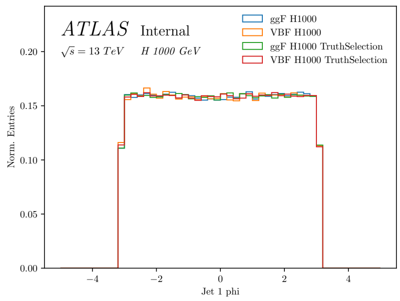
Figures E.19-E.20 shows the distributions of the 4-momentum of the input jets selected both at the reco and truth level for the pair of benchmark signals, ggF anf VBF Higgs signals with 1000 GeV mass. Figure E.21 shows the comparison of the output RNN scores obtained when the inputs jets are selected at the truth level and a comparison is done what obtained at the reco level.



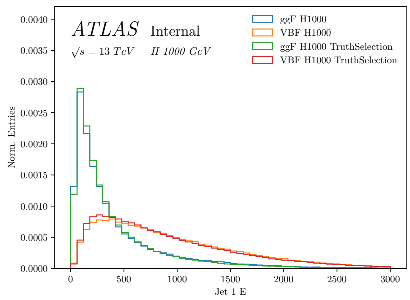
(a)



(b)

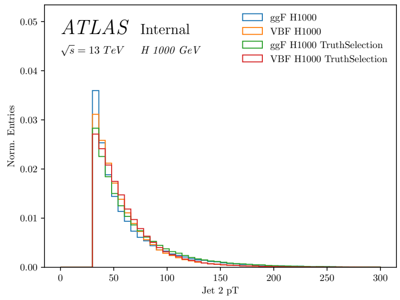


(c)

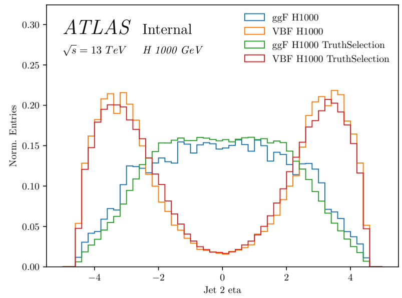


(d)

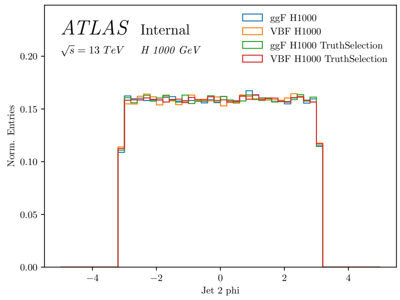
Figure E.19: Distributions of the input kinematics variables of the Jet1 comparing the jets at the truth and reco level.



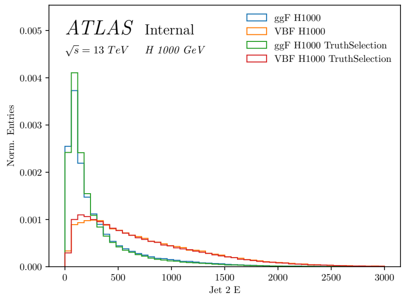
(a)



(b)

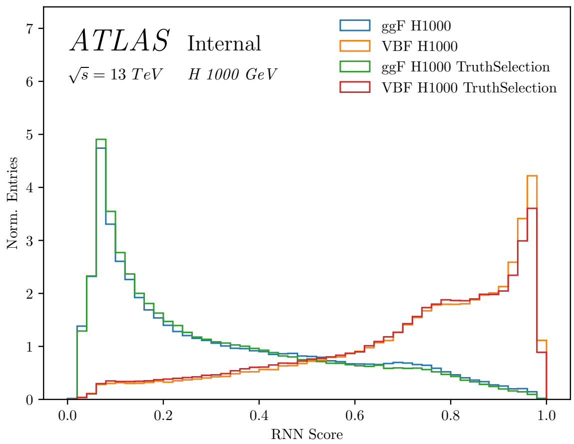


(c)



(d)

Figure E.20: Distributions of the input kinematics variables of the Jet2 comparing the jets at the truth and reco level.

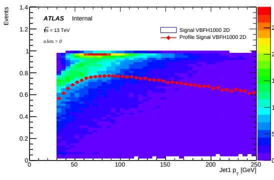


(a)

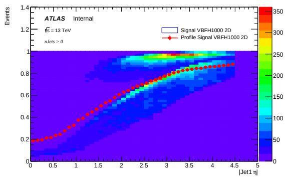
Figure E.21: Distributions of the RNN score comparing the jets at the truth and reco level.

E.6 RNN functional shape

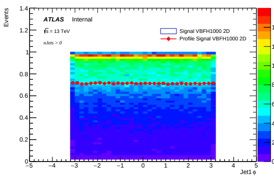
Figures E.22-E.23 show the bi-dimensional distributions of the RNN scores versus the input variables for the $VBFHiggs1000\ GeV$ signal; the functional shape is superimposed to show the trend of the RNN score as a function of the inputs.



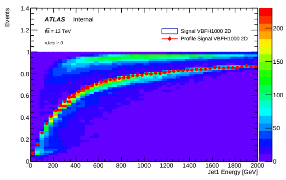
(a)



(b)

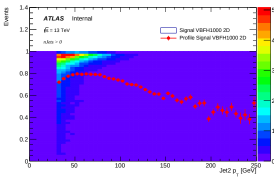


(c)

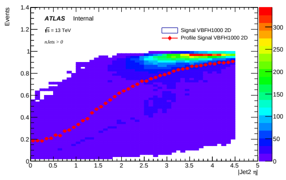


(d)

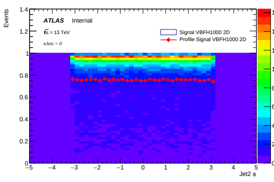
Figure E.22: Functional shape of the RNN model as function of the input variable of the first jet.



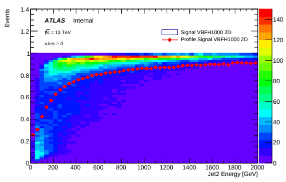
(a)



(b)



(c)



(d)

Figure E.23: Functional shape of the RNN model as function of the input variable of the second jet.

Appendix F

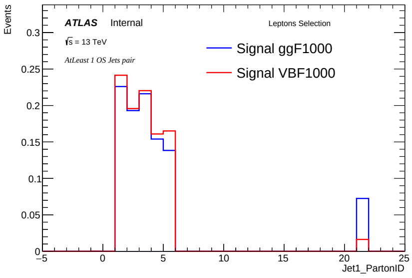
Appendix: Future perspectives of ML in Physics Analysis

New ideas and possible improvements to the ML applications showed in this thesis are presented here. Possible ideas are the improvement of the RNN model used for the VBF/ggF classification model, the adding of the Q/G informations to the model for the VBF/ggF classification, the introduction of RNN approach for the non-resonant interpretation and the Q/G jets identification in the forward region of the ATLAS detector.

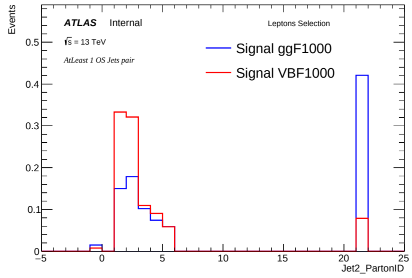
F.1 VBF/ggF classification: DNN performance adding the Quark/Gluon information

An event with the VBF production mode is expected to have 2 more jets in the topology of the final states according to production mechanism. Furthermore, these 2 additional jets are Quark-initiated jets. In principle, the features of the jets related to the flavor of the parton that initiated the jets could help further in the VBF/ggF classification. In Figure F.1 the distributions of the $PartonTruthLabelID$ for the first 6 small jets of the event are reported for both VBF and ggF 1000 GeV samples. Values of the $PartonTruthLabelID$ in range $[1, 5]$ are referred to quark-initiated jets, values equal to 21 corresponds to gluon-initiated jets and, finally, values equal to -1 referred to undefined jets.

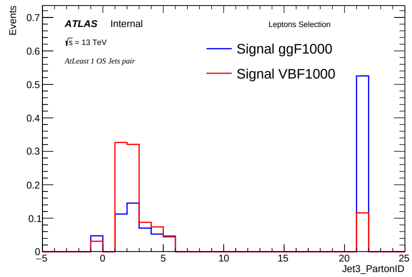
The fractions of the different kind of truth jets are evaluated and reported in table F.1. Especially the jets number $2,\ 3,\ 4$ have the biggest difference in term of the quark and gluon fraction; the VBF events have a quark jets fraction that is around two times the quark jets fraction of the ggF sample.



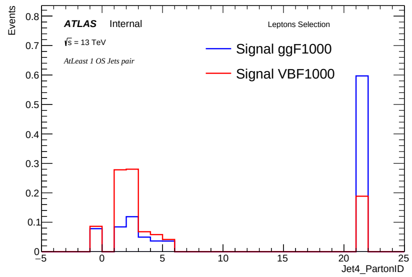
(a) *Jet 1*



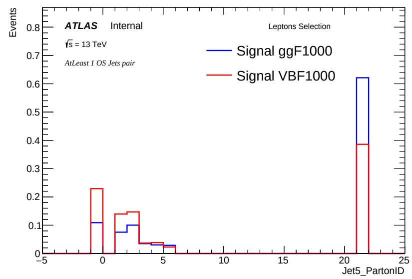
(b) *Jet 2*



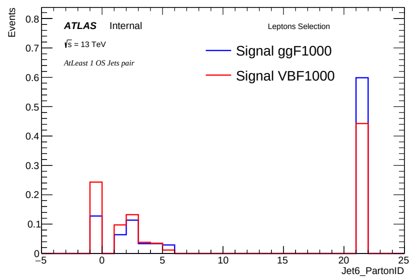
(c) *Jet 3*



(d) *Jet 4*



(e) *Jet 5*



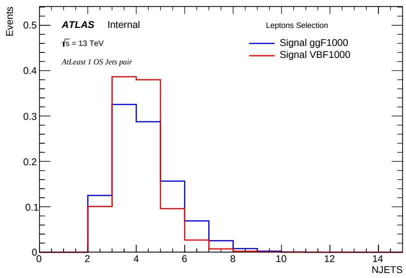
(f) *Jet 6*

Figure F.1: Shapes comparison of the Parton Truth Label ID of the 6 small jets considered as DNN input.

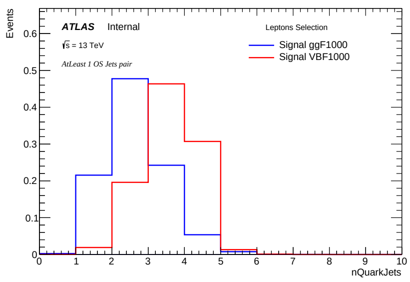
		VBF 1000 SPIN0	ggF 1000 SPIN0
Jet1	Quark	98.4%	92.7%
	Gluon	1.6%	7.3%
	Undefined	0.0%	0.0%
Jet2	Quark	91.3%	56.4%
	Gluon	7.9%	42.1%
	Undefined	0.8%	1.5%
Jet3	Quark	85.3%	42.7%
	Gluon	11.6%	52.5%
	Undefined	3.1%	4.8%
Jet4	Quark	72.6%	32.5%
	Gluon	18.8%	59.7%
	Undefined	8.6%	7.8%
Jet5	Quark	38.5%	27.0%
	Gluon	38.6%	62.1%
	Undefined	22.9%	10.9%
Jet6	Quark	31.4%	27.4%
	Gluon	44.3%	59.9%
	Undefined	24.3%	12.8%

Table F.1: Quark/Gluon fraction of the first 6 small jets used as DNN input.

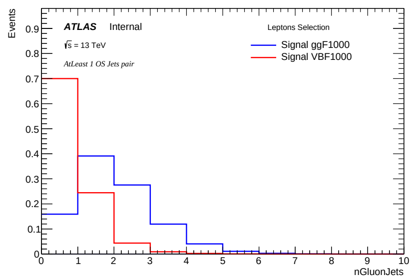
The higher quark-jets abundance in the VBF samples is supported also by the distributions of the number of truth quark jets and truth gluon jets in the VBF and ggF sample; as reported in Figure F.2 and in table F.2 the mean number of quark jets is higher for the VBF samples while the mean number of truth gluon jets is higher for the ggF sample respect to the VBF one.



(a)



(b)



(c)

Figure F.2: Shapes comparison of number of small jets (a), of the number of truth quark jets (b) and of the number of truth gluon jets (c) for the VBF and ggF signals.

	<nJets>	<nQuarkJets>	<nGluonJets>	<nUndefinedJets>
Signal SPIN0 VBF 1000	3.60	3.10	0.37	0.12
Signal SPIN0 ggF 1000	3.84	2.16	1.54	0.15

Table F.2: Mean number of jets for the VBF and ggF 1000 GeV signal mass point, divided according the truth flavor of the jets.

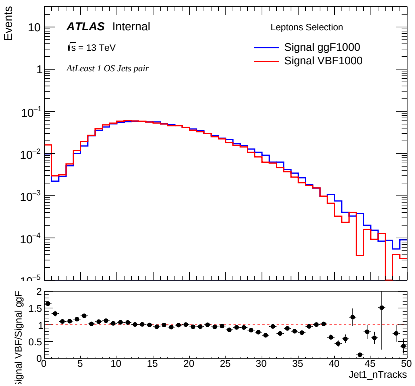
Since the quarks have a color charge respect to the gluons, the gluon-initiated jets are expected to have an higher track particles multiplicity and an higher spatial spreading. Some jets variables could help on the Quark VS Gluon jets identification; in ATLAS there is an official tagger that is based only on the number of tracks of the jets [[91]], but some variables are promising for future improvements.

In this study, 3 variables have been considered:

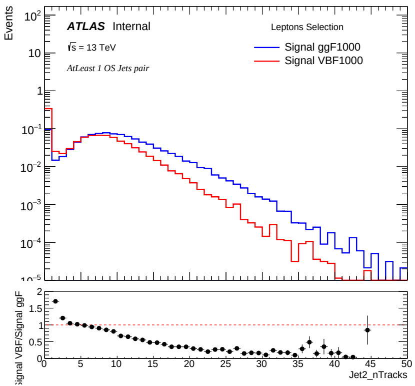
- n_{Tracks}: number of reconstructed tracks inside the jet;
- Track Width: width of the tracks of the jet;
- Calo Width: width of the calorimetric information of the jet.

The tracks inside a jets with $p_T > 500\ MeV$, primary vertex association and Loose quality for InnerDectorTrackReconstruction have been counted according to official prescription [[91]]

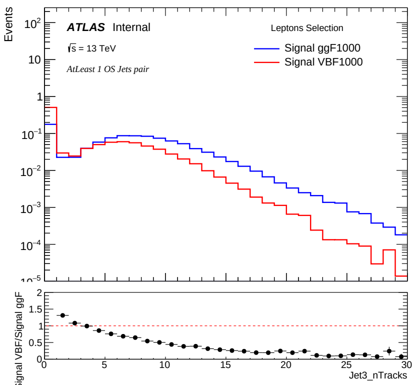
The shapes distributions for the first 6 jets are showed for n_{Tracks} in Figure F.3, for $TrackWidth$ in Figure F.4 and for $CaloWidth$ in Figure F.5.



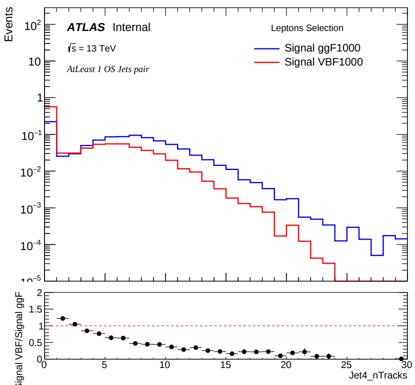
(a) *Jet 1*



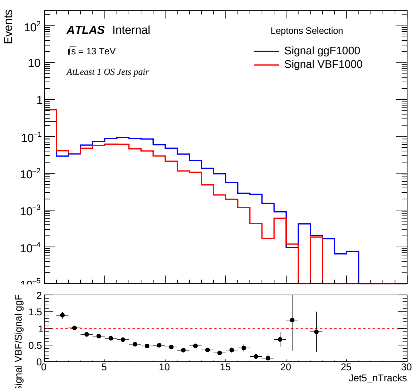
(b) *Jet 2*



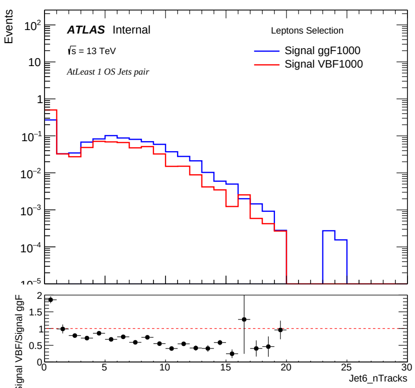
(c) *Jet 3*



(d) *Jet 4*

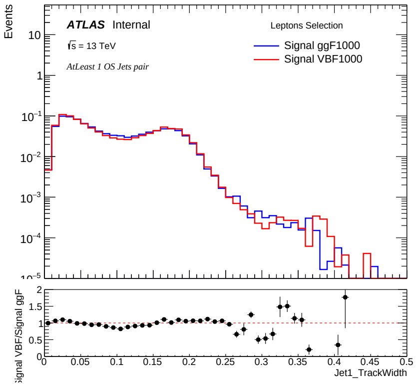


(e) *Jet 5*

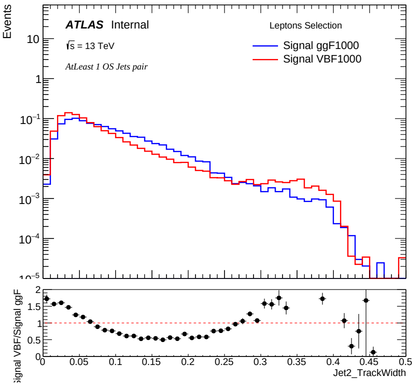


(f) *Jet 6*

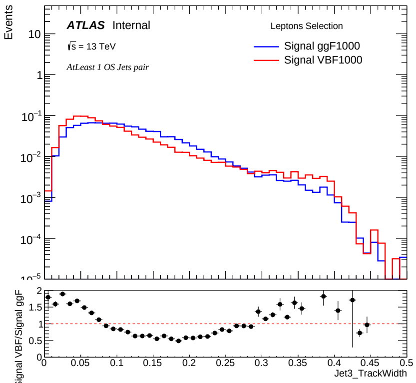
Figure F.3: Shapes comparison of the number of tracks inside the first 6 small jets for the VBF and ggF signals.



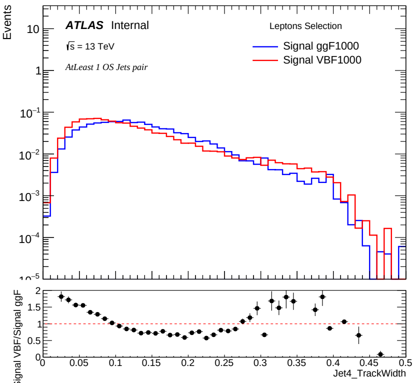
(a) *Jet 1*



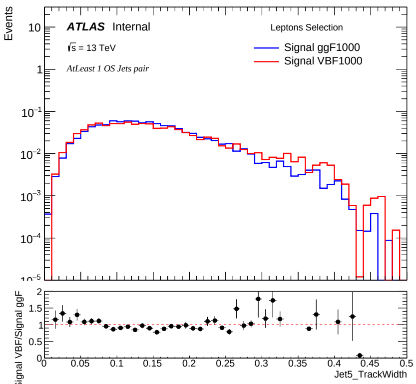
(b) *Jet 2*



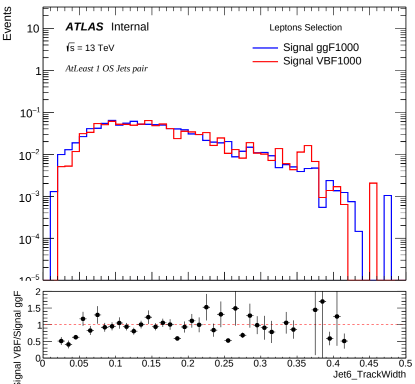
(c) *Jet 3*



(d) *Jet 4*

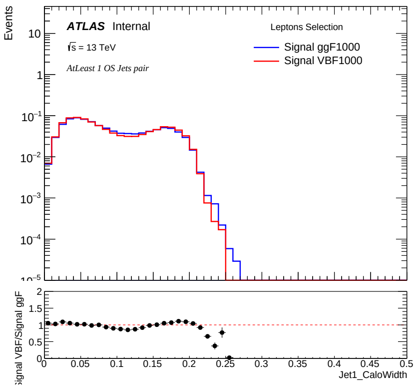


(e) *Jet 5*

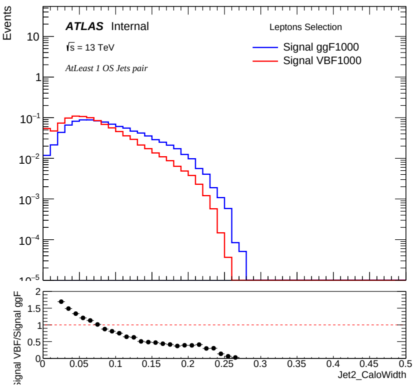


(f) *Jet 6*

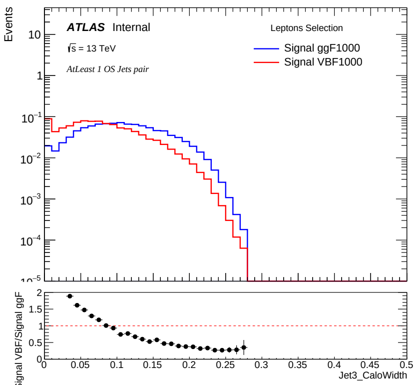
Figure F.4: Shapes comparison of the tracks width of the first 6 small jets for the VBF and ggF signals.



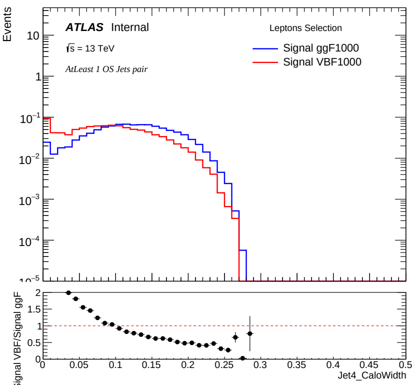
(a) *Jet 1*



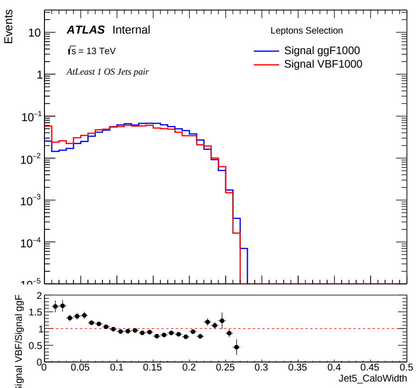
(b) *Jet 2*



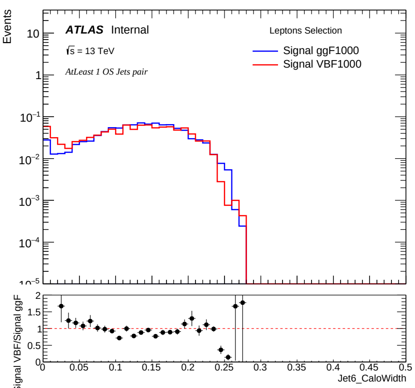
(c) *Jet 3*



(d) *Jet 4*



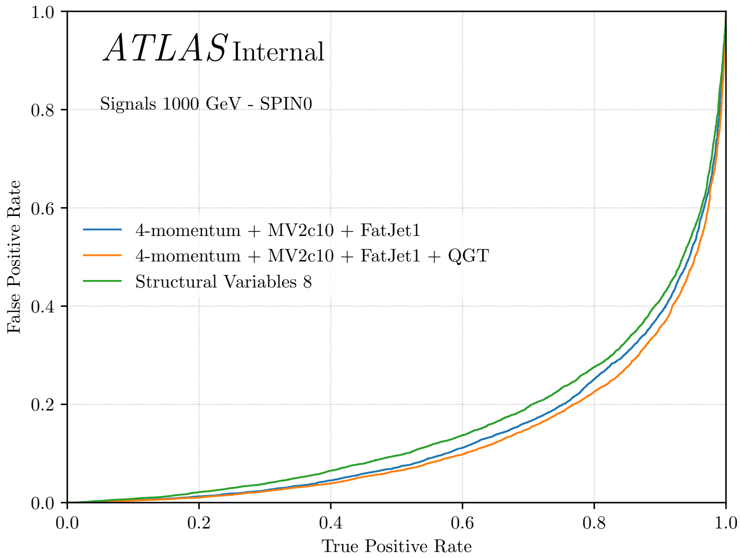
(e) *Jet 5*



(f) *Jet 6*

Figure F.5: Shapes comparison of the calorimetric width of the first 6 small jets for the VBF and ggF signals.

The DNN has been trained using the 4-momentum set of variables (low-level variables), as showed in the previous section, and adding also the Q/G variables for the first 6 small jets. The performance of the net are showed and compared in Figure F.6.



(a)

Figure F.6: ROC curve comparison using the Structural Variables, the 4-momentum and the 4-momentum with the Q/G variables.

The Q/G variables gives an improvement in the DNN VBF/ggF identification since the higher quark-jets fractions in the VBF sample. The improvement is quantified in the tables F.3 and F.4. We gain a further $10 - 15\%$ on the false positive rate decrease adding the Q/G variables to the set of 4-momentum variables.

The current WP of the cut based categorization has a false positive rate of 13% and a true positive rate of 51%; if we fix the false positive rate, the true positive rate increases up to 63% (gain of 25%) if we switch to the DNN classification based on the 4-momentum variables and it become of 66% (gain of 31%) if we add also the Q/G variables.

	False Positive Rate		
True Positive Rate	6jets 4-momentum + MV2c10 + FatJet1 4-momentum	6jets 4-momentum + MV2c10 + FatJet1 4-momentum + Quark/Gluon Variables	False Positive decrease
10%	0.5%	0.4%	25%
20%	1.3%	1.0%	30%
30%	2.5%	2.3%	9%
40%	4.5%	3.9%	15%
50%	7.1%	6.4%	11%
60%	11.2%	9.8%	14%
70%	16.4%	15%	9%
80%	25.1%	23%	9%
90%	38.5%	36%	7%

Table F.3: False positive rate for different True Positive Rate WP with and without the QG variables; the last column represent the gain in terms of the False Positive decrease.

	True Positive Rate				
False Positive Rate	maps	6jets 4-momentum + MV2c10 + FatJet1 4-momentum	True Positive improvement	6jets 4-momentum + MV2c10 + FatJet1 4-momentum + Quark/Gluon Variables	True Positive improvement
5%	25%	42%	67%	44%	76%
10%	44%	57%	30%	60%	36%
12%	49%	62%	26%	65%	32%
13%	51%	63%	25%	66%	31%
15%	54%	68%	25%	70%	29%
20%	63%	75%	19%	77%	22%
25%	70%	80%	14%	83%	18%
30%	75%	85%	13%	87%	16%
40%	83%	91%	9%	92%	11%
50%	89%	94%	6%	95%	7%
60%	94%	97%	3%	97%	3%
70%	96%	98%	2%	99%	3%
80%	98%	99%	1%	99%	1%
90%	99%	100%	1%	100%	1%

Table F.4: True positive rate for different False Positive Rate WP with the 2-dimensional approach and with/without the QG variables for the DNN approach.

F.2 RNN training: mixing leptons channels and masses values

The default RNN model used in the analysis has been trained using the Spin-0 *Higgs* VBF and ggF signals at 1000 GeV of the 2 leptons channel (using at maximum 2 jets selecting with the new overlap removal procedure described in Section 5.6.1).

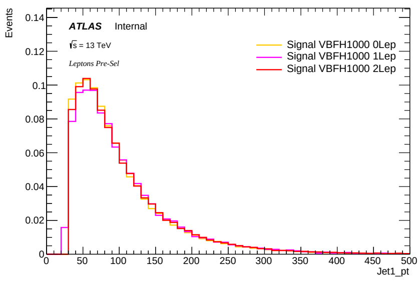
Different configurations of RNN training have been performed:

- 1) Using a mixed set of signal events from 0/1/2 leptons channel with the signal mass of 1000 GeV;
- 2) Using a mixed set of signal events from different signal mass values in the 2 leptons channel.

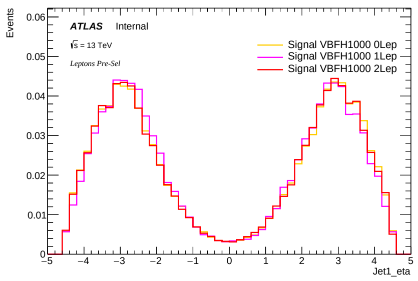
In summary, for 1), no visible difference on the performance of RNN trained using 2-lepton channel only compared to the one trained using combined 0/1/2 leptons-channel; for 2), some potential improvement can be obtained for the high mass region ($m_{VV} > 2$ TeV) by using a mixed set of signal events from different signal mass values, but given that in the high mass region the background contribution is much smaller, the gain on signal sensitivity would be not significant. Therefore, it is decided to stick to the nominal RNN setup, i.e., using only signal events from the 2-lepton channel with the signal mass of 1000 GeV.

F.2.1 Training using all three lepton channels

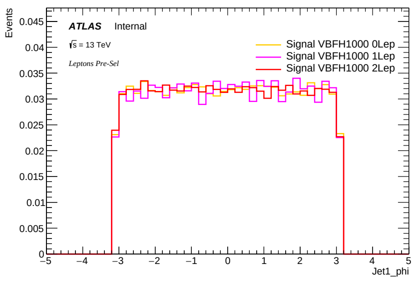
Figures F.7-F.8-F.9-F.10 show comparison of distributions of the 4-momentum component of the RNN input jets for the 3 different leptons channels. For both VBF signals and the ggF signals at a given mass point, no significant difference among the 3 lepton channels is observed.



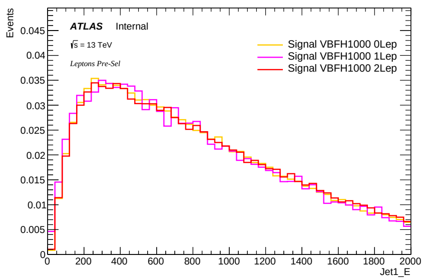
(a)



(b)

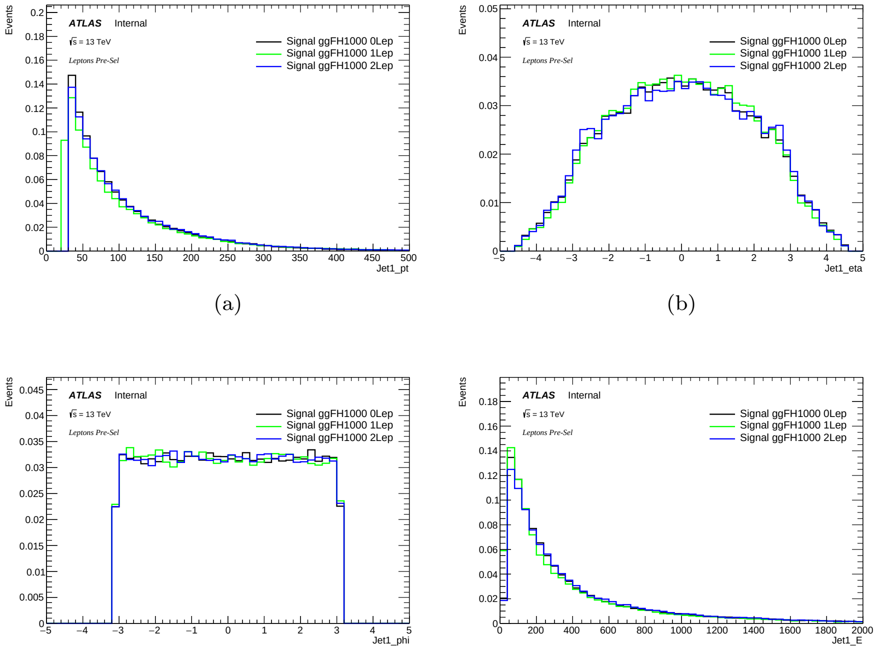


(c)



(d)

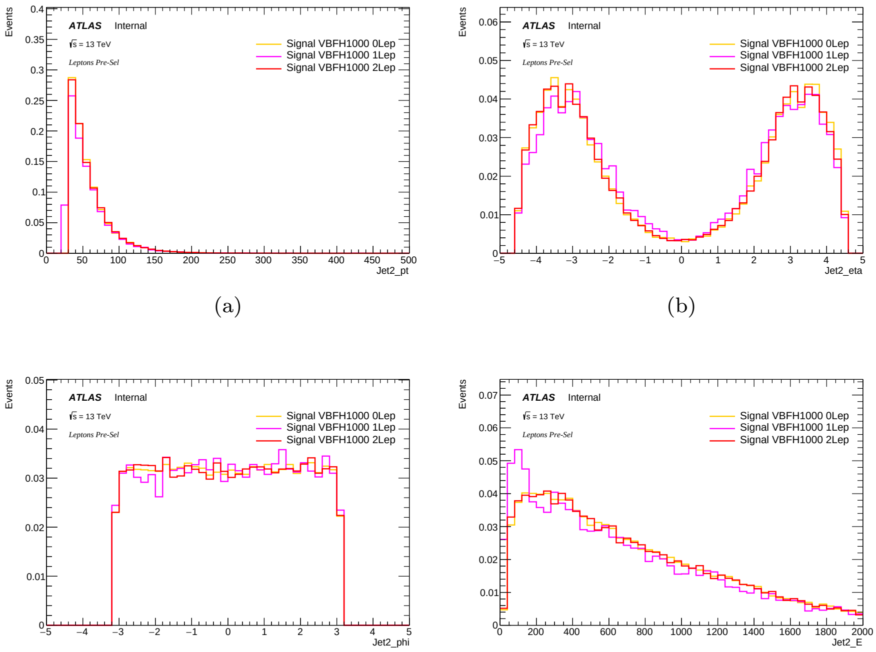
Figure F.7: Shapes distributions of the RNN input variables comparing the 3 leptons channel for the VBF signals.



(a) (b)

(c) (d)

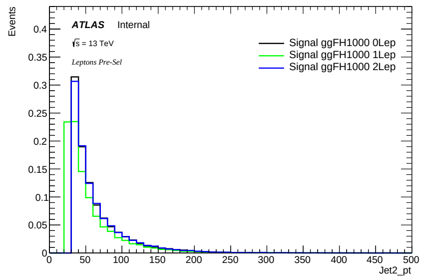
Figure F.8: Shapes distributions of the RNN input variables comparing the 3 leptons channel for the ggF signals.



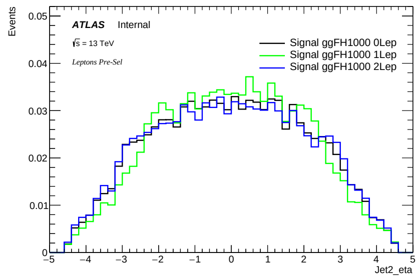
(a) (b)

(c) (d)

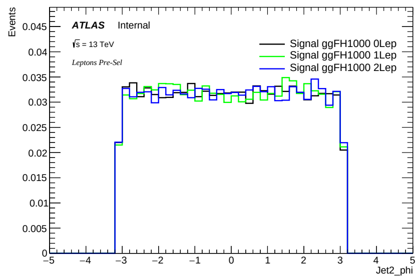
Figure F.9: Shapes distributions of the RNN input variables comparing the 3 leptons channel for the VBF signals.



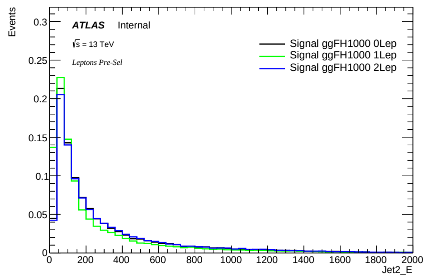
(a)



(b)



(c)



(d)

Figure F.10: Shapes distributions of the RNN input variables comparing the 3 leptons channel for the ggF signals.

A mixed set of events from 3 lepton channels has been selected for RNN training. For this test only mc16a simulations have been considered and a loose pre-selection has been performed on the samples:

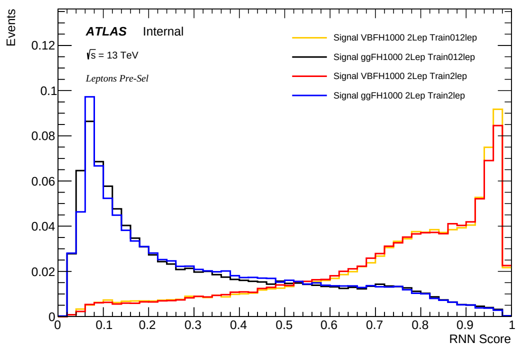
- 2 leptons: events after the dileptons selection (as usual);
- 1 lepton: events after $Met > 60\ GeV$ and $p_T(l\nu) > 75\ GeV$;
- 0 lepton: events after $Met > 250\ GeV$.

The number of events used for the training/test phase are reported in Table F.5.

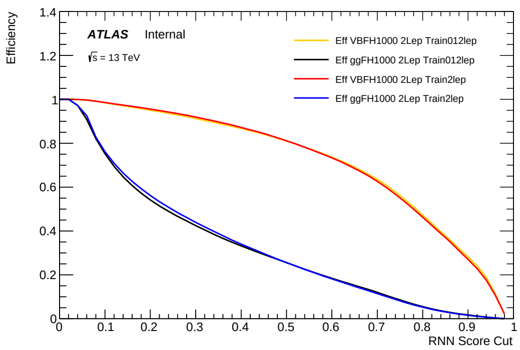
Table F.5: Events used for training/test.

	$VBF\ H$ 1000 GeV	$ggF\ H$ 1000 GeV
0 lep	116410	93217
1 lep	61623	46402
2 lep	88641	66057

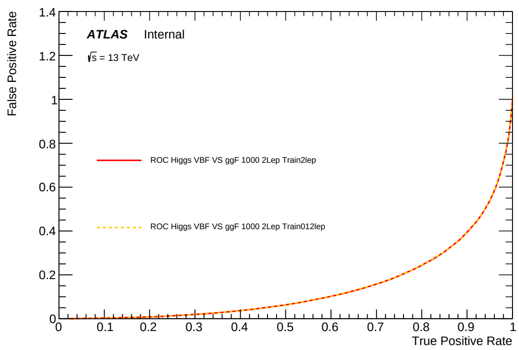
The mixed trained model is compared to the default 2-lepton-only model. Figures F.11-F.12-F.13 show comparison of distributions of the RNN score, the signal selection efficiencies and ROC curves, evaluated for the 2-lepton, 1-lepton and 0-lepton signal events, respectively. The performance of the mixed model is consistent with the default model, and no obvious difference is expected in the analysis.



(a)

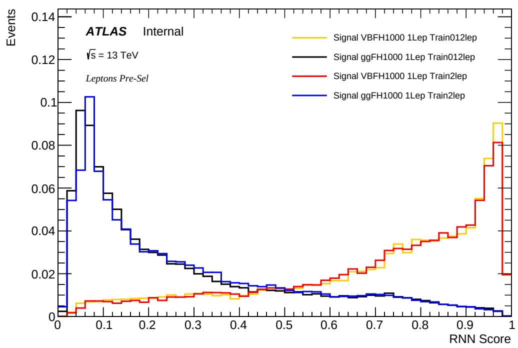


(b)

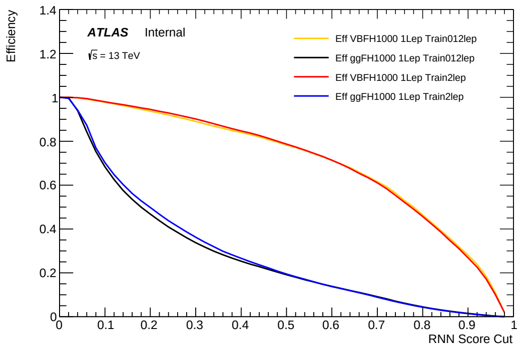


(c)

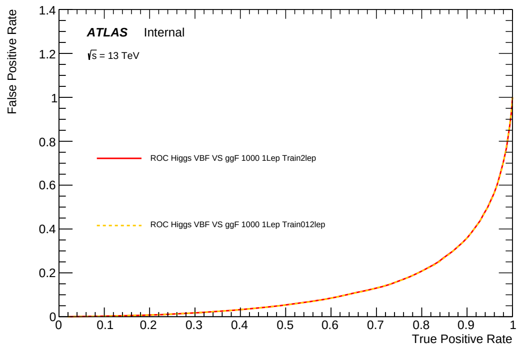
Figure F.11: Performances comparison of the 2 lepton model VS the mixed model for the 2 leptons signals.



(a)

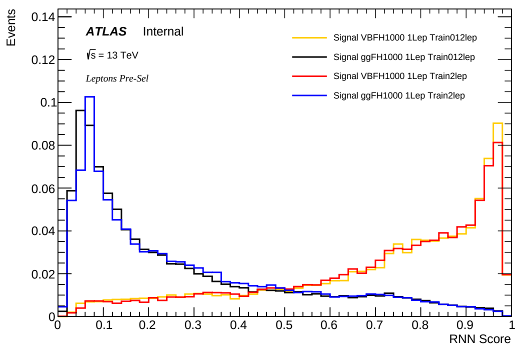


(b)

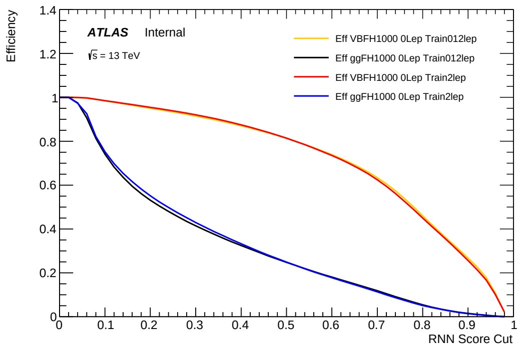


(c)

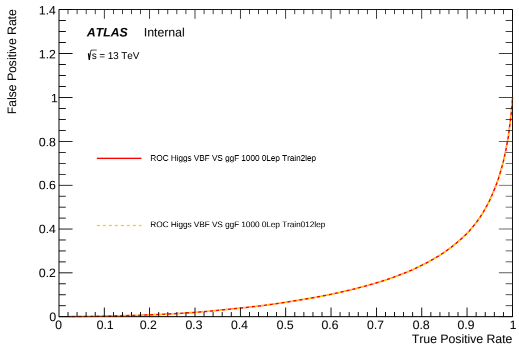
Figure F.12: Performances comparison of the 2 lepton model VS the mixed model for the 1 leptons signals.



(a)



(b)

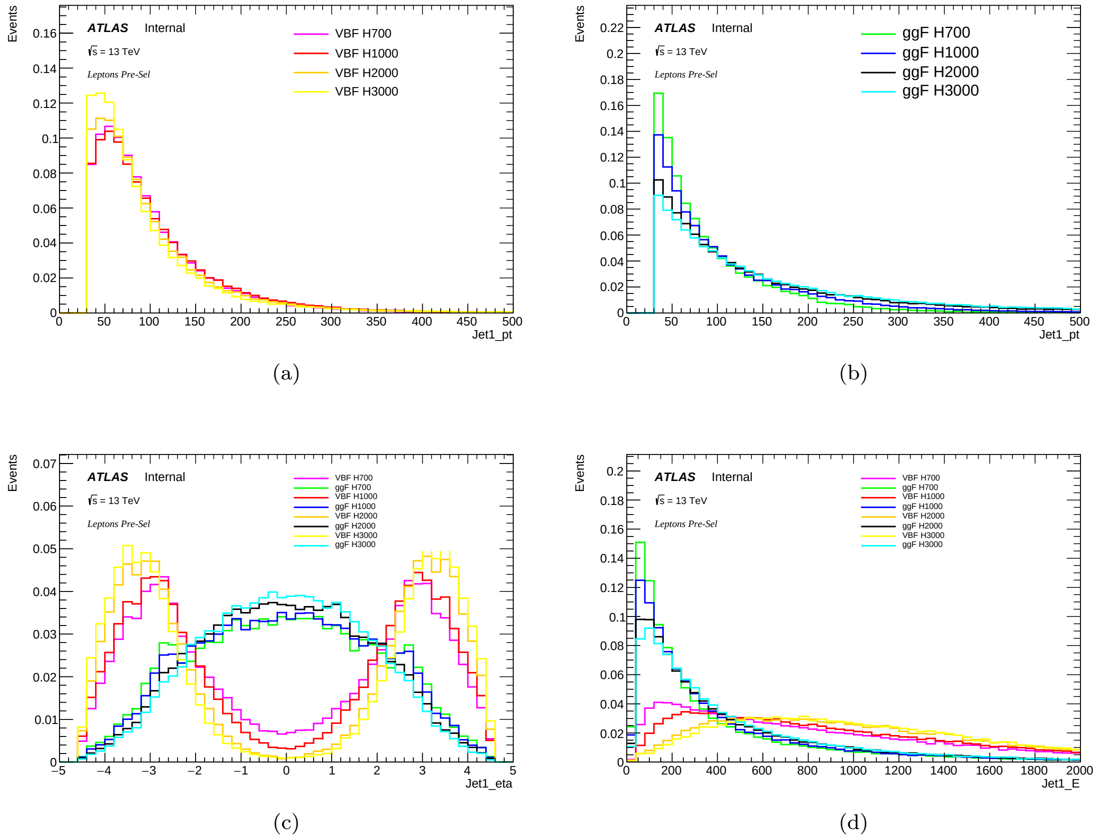


(c)

Figure F.13: Performances comparison of the 2 lepton model VS the mixed model for the 0 leptons signals.

F.2.2 Training using mixed signal mass points

A mixed set of signal events with different mass points has been considered for training a RNN model, and it is compared to the default model which is trained using H 1000 GeV only. Figures F.14-F.15 show the comparison of distributions of the input variables. Some small difference is present among the different mass hypotheses.



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

Figure F.14: Comparison of the p_T, η and energy distribution of the first jet of different mass signals of the 2 leptons channel.

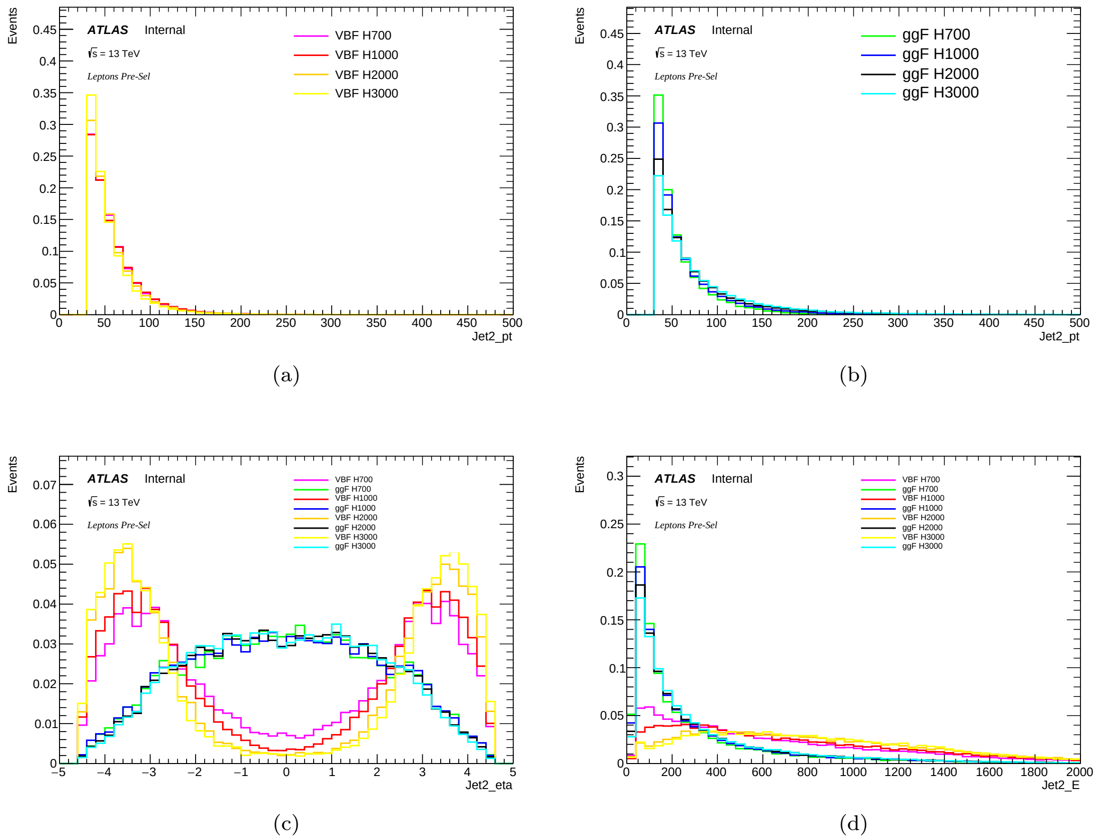


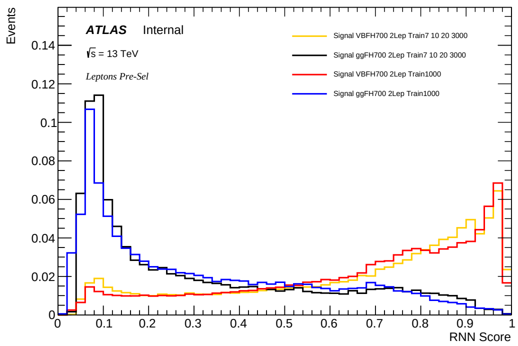
Figure F.15: Comparison of the p_T, η and energy distribution of the second jet of different mass signals of the 2 leptons channel.

The number of events used for training/test phase are reported in the Table F.6.

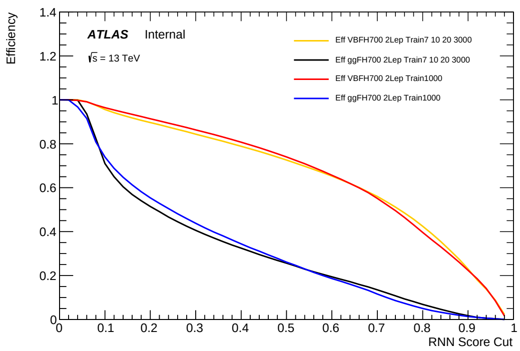
Table F.6: Events used for training/test.

	VBF	ggF
700 GeV	70789	56341
1000 GeV	76009	66057
2000 GeV	75283	76072
3000 GeV	67676	68537

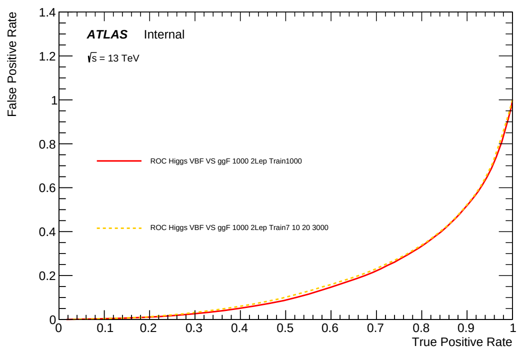
The mixed mass trained model is compared to the default $H1000\ GeV$-only model. Figures F.16-F.17-F.18-F.19 show distributions of the RNN score, the signal selection efficiencies and ROC curves for signal events with various mass points. The performance of the mixed model is similar to the default model at low mass, but it becomes to perform better at high mass range. The mixed model could improve further the RNN performances and also a parameterised-RNN model could be investigated.



(a)

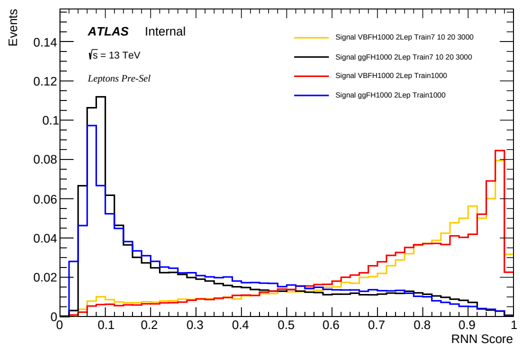


(b)

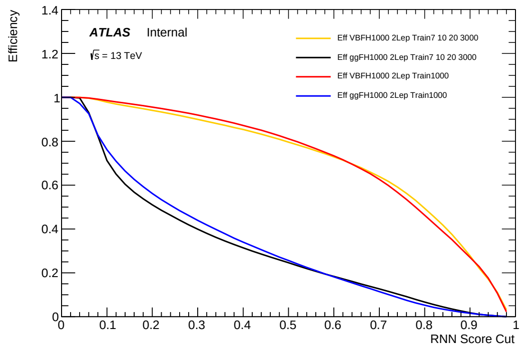


(c)

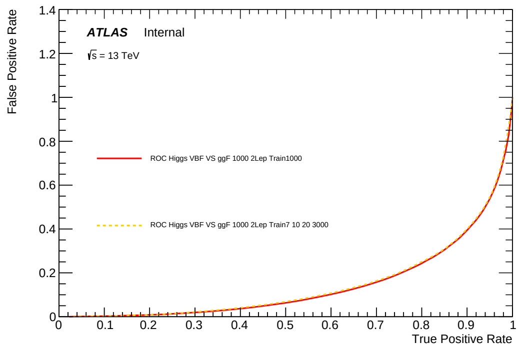
Figure F.16: Performances comparison of the H 1000 GeV model VS the mixed mass model for the 700 GeV signals.



(a)

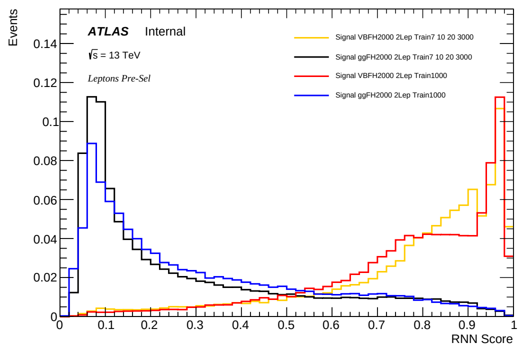


(b)

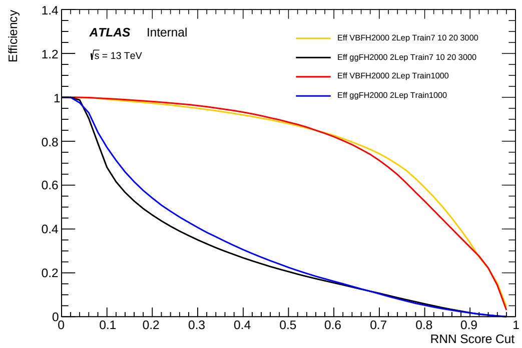


(c)

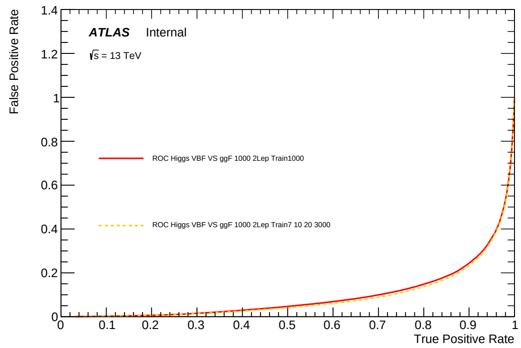
Figure F.17: Performances comparison of the H 1000 GeV model VS the mixed mass model for the 1000 GeV signals.



(a)

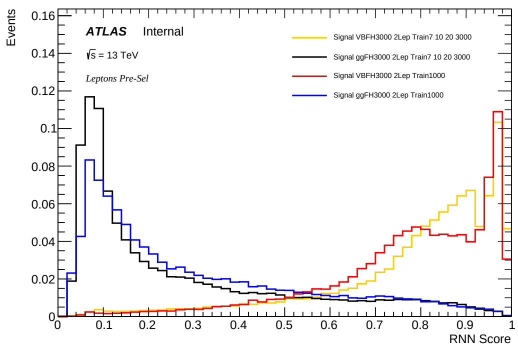


(b)

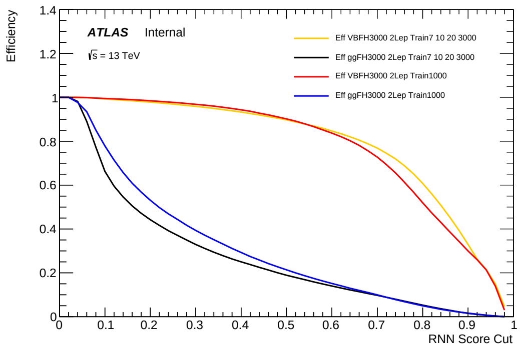


(c)

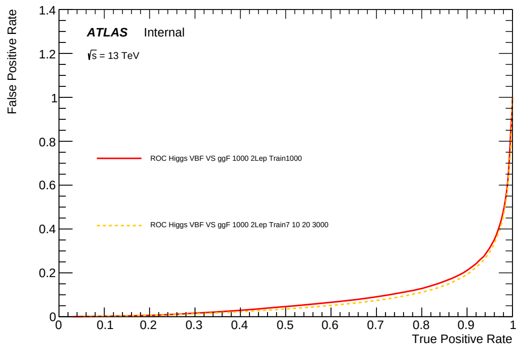
Figure F.18: Performances comparison of the H 1000 GeV model VS the mixed mass model for the 2000 GeV signals.



(a)



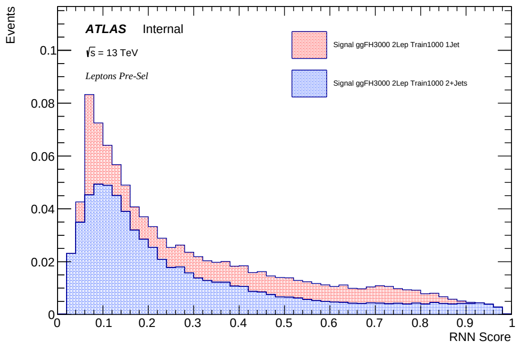
(b)



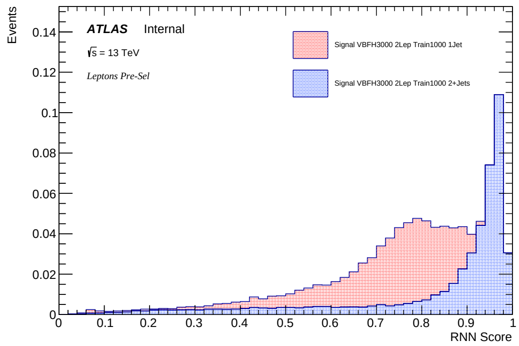
(c)

Figure F.19: Performances comparison of the H 1000 GeV model VS the mixed mass model for the 3000 GeV signals.

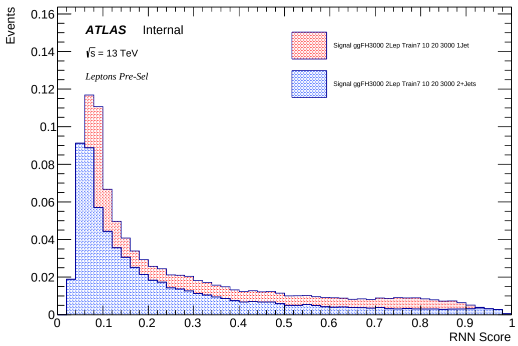
The contributions of the $1Jet$ and $\geq 2Jets$ events to the RNN score shape have been investigated and showed in Figure F.20 for the 3000 GeV signal. The RNN is trained using the mixed model.



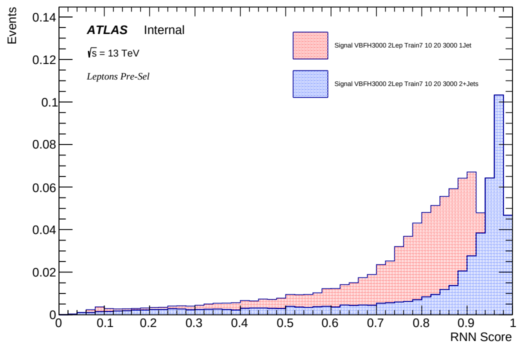
(a)



(b)



(c)



(d)

Figure F.20: Stack plot of the RNN score; the $1 - Jet$ and the $\geq 2Jets$ events contributions have been separated.

Figure F.21 shows the shapes comparison of the RNN score of the default model (red) with the mixed model (blue).

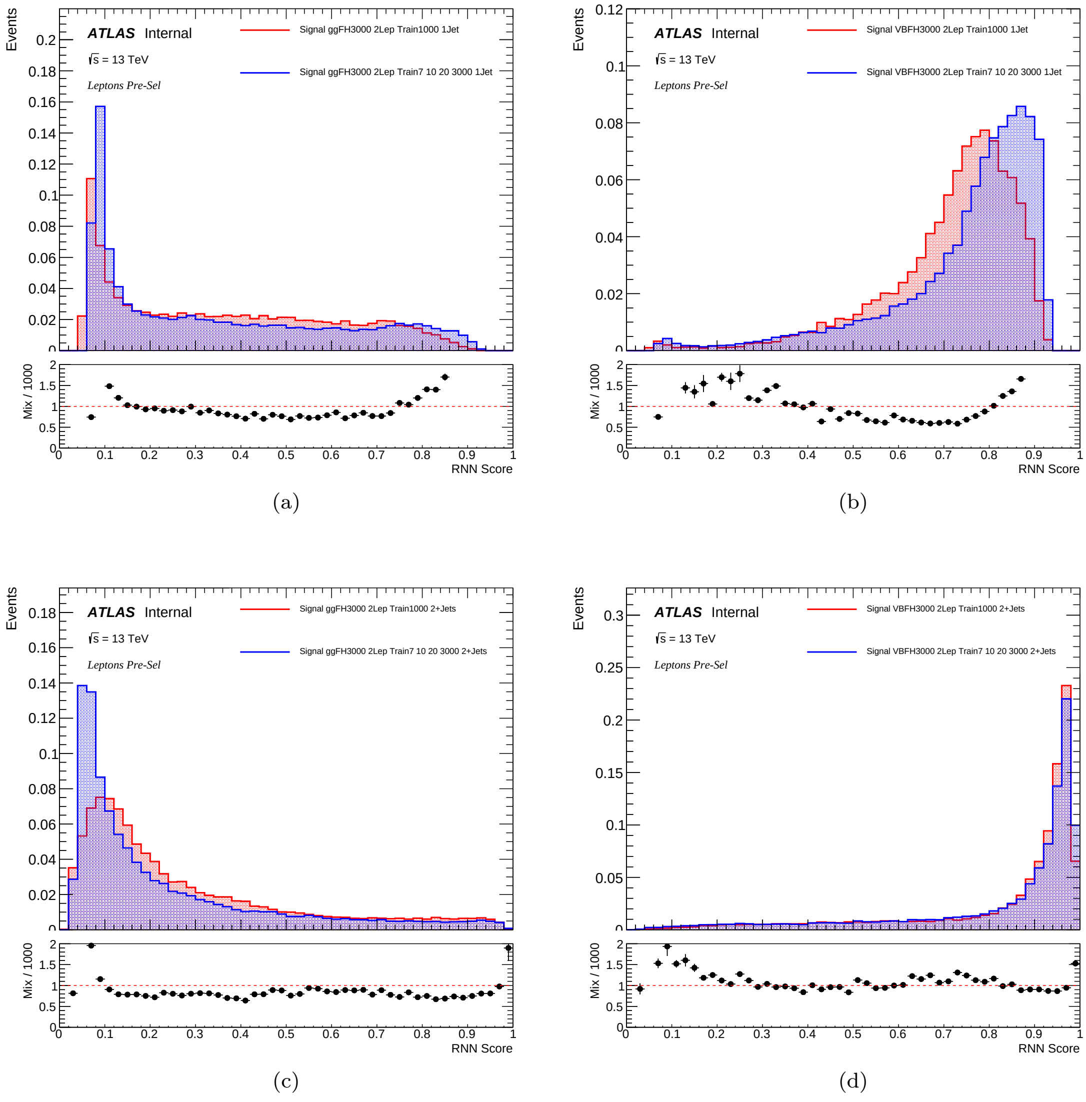
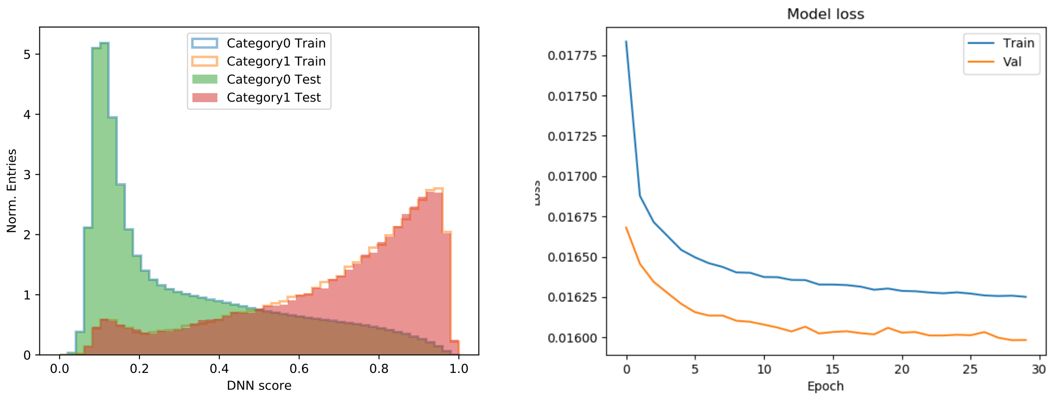


Figure F.21: Shapes plot of the RNN score; the $1 - Jet$ and the $\geq 2Jets$ events contributions have been separated.

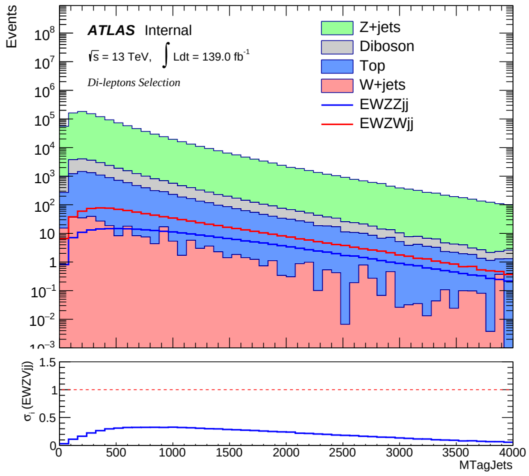
F.3 RNN for non resonant interpretation

A RNN approach is under development for the non-resonant interpretation. The aim is to categorise the EW signals (ZZ/ZW + jj) versus the SM backgrounds (Z+jets, SM Dibosons, Top). A similar architecture to the one described for the resonant interpretation 5.6 has been testes. This time the input variables are 5 (the 4-momentum + Tracks multiplicity) for 6 jets as maximum. Figure F.22 show the shapes of the score obtained for the EW ZZjj signal and for the full expected background and the loss function during the training phase to test the goodness of the model. Figures F.23 show the comparison of the invariant mass of the dijets system m_{jj}^{Tag} and of the new RNN score. The asymptotic formula was used to evaluated the discrimination power of the two variables to the EW signals; the MC used refers to the full run-II dataset. Furthermore, the ROC curves in the two cases are compared in Figure F.23-(c). The RNN score shows promising discrimination power with a significance combining all the bins of 2.7 while the significance evaluated on the m_{jj}^{Tag} reaches only 1.3.

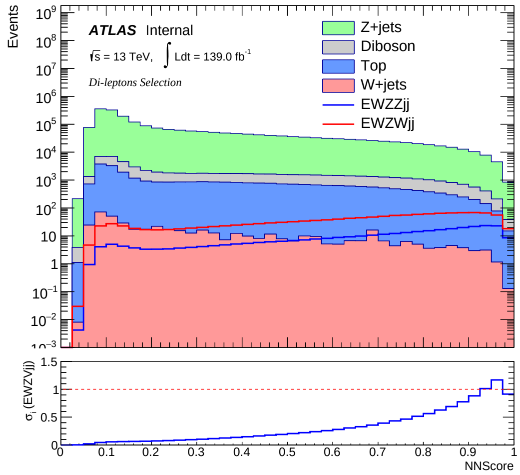


(a) (b)

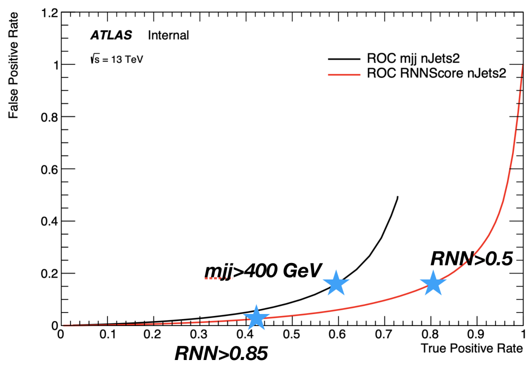
Figure F.22: RNN score (a) and loss function during the training phase (b).



(a)



(b)

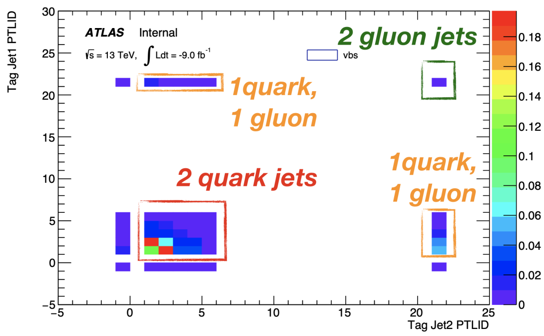


(c)

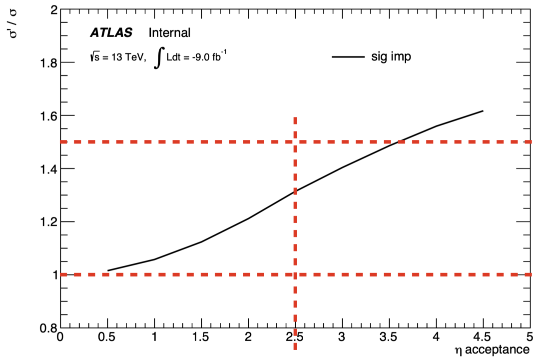
Figure F.23: Stack plot of the SM background distributions for the m_{jj}^{Tag} (a) and for the RNN score (b); the EW signals are superimposed. The bottom panel show the local significance evaluated using the asymptotic formula.

F.4 Q/G identification in forward region

In the context of the non-resonant analysis the possibility to develop some Quark/Gluon jets identification in the forward region of the detector has been studied. Figure F.24-(a) shows the truth information of the parton associated to the tag jets pair selected; if the gluon jets could be removed it is expected to reduce a lot of background, as discussed in section 5.11.5. Figure F.24-(b) shows the improvement on the expected significance if events that have not two quark-jets as tag jets are removed as a function of the η range in which the removal is performed. An improvement up to 60% is expected if the acceptance range of the tagging would be extended up to $\eta = 4.5$. This motivates the possibility to use calorimetric informations in the forward region of the detector, where there is not the tracker, to can identify quark/gluon-jets.



(a)



(b)

Figure F.24: Parton Truth Label ID for the tag jets pair for a EW signal sample (a) and improvement on the expected significance if the Q/G tagging range would extend up to the entire acceptance of the detector (b).

Bibliography

[1] Steven Weinberg. “The Making of the Standard Model”. In: *Eur. Phys. J. C* 34.hep-ph/0401010 (2003), 5–13. 21 p. ; streaming video. URL: `https://cds.cern.ch/record/799984`.

[2] Murray Gell-Mann. “A Schematic Model of Baryons and Mesons”. In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: `10.1016/S0031-9163(64)92001-3`.

[3] G Zweig. “An SU_3 model for strong interaction symmetry and its breaking; Version 2”. In: CERN-TH-412 (1964), 80 p. URL: `https://cds.cern.ch/record/570209`.

[4] E. D. Bloom et al. “High-Energy Inelastic $e - p$ Scattering at 6° and 10°”. In: *Phys. Rev. Lett.* 23 (16 1969), pp. 930–934. DOI: `10.1103/PhysRevLett.23.930`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.23.930`.

[5] M. Breidenbach et al. “Observed Behavior of Highly Inelastic Electron-Proton Scattering”. In: *Phys. Rev. Lett.* 23 (16 1969), pp. 935–939. DOI: `10.1103/PhysRevLett.23.935`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.23.935`.

[6] C. N. Yang and R. L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. In: *Phys. Rev.* 96 (1 1954), pp. 191–195. DOI: `10.1103/PhysRev.96.191`. URL: `https://link.aps.org/doi/10.1103/PhysRev.96.191`.

[7] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (1964). [,160(1964)], pp. 508–509. DOI: `10.1103/PhysRevLett.13.508`.

[8] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (1964). [,157(1964)], pp. 321–323. DOI: `10.1103/PhysRevLett.13.321`.

[9] H. Weyl. “Theory of Groups and Quantum Mechanics.” In: (1929).

[10] S. L. Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: `10.1016/0029-5582(61)90469-2`.

[11] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: `10.1103/PhysRevLett.19.1264`.

[12] Abdus Salam. “Weak and Electromagnetic Interactions”. In: *Conf. Proc.* C680519 (1968), pp. 367–377. DOI: `10.1142/9789812795915_0034`.

[13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (1964). [,162(1964)], pp. 585–587. DOI: `10.1103/PhysRevLett.13.585`.

[14] T. W. B. Kibble. “Spontaneous symmetry breaking in gauge theories”. In: *Phil. Trans. Roy. Soc. Lond.* A373.2032 (2014), p. 20140033. DOI: `10.1098/rsta.2014.0033`.

[15] Laura Reina. “Higgs + EW Theory. 2019 CERN-Fermilab HCP Summer School”. In: (2019). URL: `https://cds.cern.ch/record/2688157`.

[16] M. Aaboud et al. “Measurement of the Higgs boson mass in the $H \to ZZ \to 4l$ and $H \to \gamma\gamma$ channels with s=13 TeV pp collisions using the ATLAS detector”. In: *Physics Letters B* 784 (2018), 345?366. ISSN: 0370-2693. DOI: `10.1016/j.physletb.2018.07.050`. URL: `http://dx.doi.org/10.1016/j.physletb.2018.07.050`.

[17] M. Aaboud et al. "Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ with the ATLAS detector". In: *The European Physical Journal C* 77.6 (2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4900-z. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4900-z.

[18] David J. Gross and Frank Wilczek. "Asymptotically Free Gauge Theories. I". In: *Phys. Rev. D* 8 (10 1973), pp. 3633–3652. DOI: 10.1103/PhysRevD.8.3633. URL: https://link.aps.org/doi/10.1103/PhysRevD.8.3633.

[19] H David Politzer. "Asymptotic freedom: An approach to strong interactions". In: *Physics Reports* 14.4 (1974), pp. 129 –180. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(74)90014-3. URL: http://www.sciencedirect.com/science/article/pii/0370157374900143.

[20] Radja Boughezal et al. "Z-Boson Production in Association with a Jet at Next-To-Next-To-Leading Order in Perturbative QCD". In: *Physical Review Letters* 116.15 (2016). ISSN: 1079-7114. DOI: 10.1103/physrevlett.116.152001. URL: http://dx.doi.org/10.1103/PhysRevLett.116.152001.

[21] Radja Boughezal, Xiaohui Liu, and Frank Petriello. "Phenomenology of theZboson plus jet process at NNLO". In: *Physical Review D* 94.7 (2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.074015. URL: http://dx.doi.org/10.1103/PhysRevD.94.074015.

[22] M. Aaboud et al. "Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$TeV with the ATLAS detector". In: *The European Physical Journal C* 79.6 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7027-6. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7027-6.

[23] G. Aad et al. "Measurement of the ZZ Production Cross Section in pp Collisions at $\sqrt{s} =$ 13TeV with the ATLAS Detector". In: *Physical Review Letters* 116.10 (2016). ISSN: 1079-7114. DOI: 10.1103/physrevlett116.101801. URL: http://dx.doi.org/10.1103/PhysRevLett.116.101801.

[24] Massimiliano Grazzini, Stefan Kallweit, and Dirk Rathlev. "ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD". In: *Physics Letters B* 750 (2015), pp. 407 –410. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2015.09.055. URL: http://www.sciencedirect.com/science/article/pii/S0370269315007303.

[25] Fabrizio Caola et al. "QCD corrections to ZZ production in gluon fusion at the LHC". In: *Phys. Rev. D* 92 (9 2015), p. 094028. DOI: 10.1103/PhysRevD.92.094028. URL: https://link.aps.org/doi/10.1103/PhysRevD.92.094028.

[26] G. Aad et al. "Measurements of the Total and Differential Higgs Boson Production Cross Sections Combining the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ Decay Channels at $\sqrt{s} = 8TeV$ with the ATLAS Detector". In: *Physical Review Letters* 115.9 (2015). ISSN: 1079-7114. DOI: 10.1103/physrevlett.115.091801. URL: http://dx.doi.org/10.1103/PhysRevLett.115.091801.

[27] G. Aad et al. "Search for an additional, heavy Higgs boson in the $H \to ZZ$ decay channel at $\sqrt{s} = 8$ TeV in *pp* collision data with the ATLAS detector". In: *The European Physical Journal C* 76.1 (2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3820-z. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3820-z.

[28] U. Baur and D. Rainwater. "Probing neutral gauge boson self-interactions in ZZ production at hadron colliders". In: *Phys. Rev. D* 62 (11 2000), p. 113011. DOI: 10.1103/PhysRevD.62.113011. URL: https://link.aps.org/doi/10.1103/PhysRevD.62.113011.

[29] *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment.* Tech. rep. ATLAS-CONF-2019-044. Geneva: CERN, 2019. URL: http://cds.cern.ch/record/2690717.

[30] Matteo Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Physics Letters B* 710.4-5 (2012), 612?622. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.03.013. URL: http://dx.doi.org/10.1016/j.physletb.2012.03.013.

[31] Peter Barnreuther, Micha? Czakon, and Alexander Mitov. "Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Physical Review Letters* 109.13 (2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.109.132001. URL: http://dx.doi.org/10.1103/PhysRevLett.109.132001.

[32] ATLAS Collaboration. *Measurement of the $t\bar{t}$ production cross-section using e? events with b-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector*. 2014. arXiv: 1406.5375 [hep-ex].

[33] M. Aaboud et al. "Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s}$ = 13?TeV using the ATLAS detector". In: *Physical Review D* 98.1 (2018). ISSN: 2470-0029. DOI: 10.1103/physrevd.98.012003. URL: http://dx.doi.org/10.1103/PhysRevD.98.012003.

[34] G. Aad et al. "Measurements of top-quark pair differential and double-differential cross-sections in the l+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *The European Physical Journal C* 79.12 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7525-6. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7525-6.

[35] V. Khachatryan et al. "Measurement of the $\mathrm{t\bar{t}}$ production cross section in the $e\mu$ channel in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV". In: *Journal of High Energy Physics* 2016.8 (2016). ISSN: 1029-8479. DOI: 10.1007/jhep08(2016)029. URL: http://dx.doi.org/10.1007/JHEP08(2016)029.

[36] V. Khachatryan et al. "Measurement of the $\mathrm{t\bar{t}}$ production cross section using events in the eμ final state in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *The European Physical Journal C* 77.3 (2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4718-8. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4718-8.

[37] A. M. Sirunyan et al. "Measurement of the inclusive $\mathrm{t\bar{t}}$ cross section in pp collisions at $\sqrt{s} = 5.02$ TeV using final states with at least one charged lepton". In: *Journal of High Energy Physics* 2018.3 (2018). ISSN: 1029-8479. DOI: 10.1007/jhep03(2018)115. URL: http://dx.doi.org/10.1007/JHEP03(2018)115.

[38] Benjamin W. Lee, C. Quigg, and H. B. Thacker. "Strength of Weak Interactions at Very High Energies and the Higgs Boson Mass". In: *Phys. Rev. Lett.* 38 (16 1977), pp. 883–885. DOI: 10.1103/PhysRevLett.38.883. URL: https://link.aps.org/doi/10.1103/PhysRevLett.38.883.

[39] Michael S. Chanowitz and Mary K. Gaillard. "The TeV physics of strongly interacting W's and Z's". In: *Nuclear Physics B* 261 (1985), pp. 379 –431. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(85)90580-2. URL: http://www.sciencedirect.com/science/article/pii/0550321385905802.

[40] Micha? Szleper. "The Higgs boson and the physics of WW scattering before and after Higgs discovery". In: (2014). arXiv: 1412.8367 [hep-ph].

[41] N. Arkani-Hamed et al. "The Littlest Higgs". In: *JHEP* 07 (2002), p. 034. DOI: 10.1088/1126-6708/2002/07/034. arXiv: hep-ph/0206021 [hep-ph].

[42] *Observation of electroweak production of two jets in association with a Z-boson pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2019-033. Geneva: CERN, 2019. URL: http://cds.cern.ch/record/2682845.

[43] Duccio Pappadopulo et al. "Heavy vector triplets: bridging theory and data". In: *Journal of High Energy Physics* 2014.9 (2014), p. 60. ISSN: 1029-8479. DOI: 10.1007/JHEP09(2014)060. URL: https://doi.org/10.1007/JHEP09(2014)060.

[44] V. Barger, W. Y. Keung, and Ernest Ma. “Gauge model with light W and Z bosons”. In: *Phys. Rev. D* 22 (3 1980), pp. 727–737. DOI: 10.1103/PhysRevD.22.727. URL: https://link.aps.org/doi/10.1103/PhysRevD.22.727.

[45] Roberto Contino et al. “On the effect of resonances in composite Higgs phenomenology”. In: *Journal of High Energy Physics* 2011.10 (2011). ISSN: 1029-8479. DOI: 10.1007/jhep10(2011)081. URL: http://dx.doi.org/10.1007/JHEP10(2011)081.

[46] Lisa Randall and Raman Sundrum. “Large Mass Hierarchy from a Small Extra Dimension”. In: *Physical Review Letters* 83.17 (1999), 3370?3373. ISSN: 1079-7114. DOI: 10.1103/physrevlett.83.3370. URL: http://dx.doi.org/10.1103/PhysRevLett.83.3370.

[47] Kaustubh Agashe et al. “Warped gravitons at the CERN LHC and beyond”. In: *Physical Review D* 76.3 (2007). ISSN: 1550-2368. DOI: 10.1103/physrevd.76.036006. URL: http://dx.doi.org/10.1103/PhysRevD.76.036006.

[48] Walter D. Goldberger and Mark B. Wise. “Modulus Stabilization with Bulk Fields”. In: *Phys. Rev. Lett.* 83 (24 1999), pp. 4922–4925. DOI: 10.1103/PhysRevLett.83.4922. URL: https://link.aps.org/doi/10.1103/PhysRevLett.83.4922.

[49] Walter D. Goldberger and Mark B. Wise. “Phenomenology of a stabilized modulus”. In: *Physics Letters B* 475.3-4 (2000), 275?279. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(00)00099-x. URL: http://dx.doi.org/10.1016/S0370-2693(00)00099-X.

[50] Alexandra Oliveira. “Gravity particles from Warped Extra Dimensions, predictions for LHC”. In: (2014). arXiv: 1404.0102 [hep-ph].

[51] Vernon Barger and Muneyuki Ishida. “Randall?Sundrum reality at the LHC”. In: *Physics Letters B* 709.3 (2012), 185?191. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.01.073. URL: http://dx.doi.org/10.1016/j.physletb.2012.01.073.

[52] Csaba Csáki, Jay Hubisz, and Seung J. Lee. “Radion phenomenology in realistic warped space models”. In: *Phys. Rev. D* 76 (12 2007), p. 125015. DOI: 10.1103/PhysRevD.76.125015. URL: https://link.aps.org/doi/10.1103/PhysRevD.76.125015.

[53] ATLAS Collaboration. “Searches for heavy ZZ and ZW resonances in the $\ell\ell qq$ and $\nu\nu qq$ final states in pp collisions at $\sqrt{s} = 13\,TeV$ with the ATLAS detector”. In: *JHEP* 03 (2018), p. 009. DOI: 10.1007/JHEP03(2018)009. arXiv: 1708.09638 [hep-ex].

[54] “ATLAS central solenoid: Technical design report”. In: (1997).

[55] “ATLAS barrel toroid: Technical design report”. In: (1997).

[56] “ATLAS endcap toroids: Technical design report”. In: (1997).

[57] Georges Aad et al. “Performance of the ATLAS Trigger System in 2010”. In: *Eur. Phys. J.* C72 (2012), p. 1849. DOI: 10.1140/epjc/s10052-011-1849-1. arXiv: 1110.1530 [hep-ex].

[58] G. Aad et al. “Performance of the ATLAS Trigger System in 2010”. In: *The European Physical Journal C* 72.1 (2012). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1849-1. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1849-1.

[59] “The ATLAS Data Acquisition and High Level Trigger system”. In: *Journal of Instrumentation* 11.06 (2016), P06008–P06008. DOI: 10.1088/1748-0221/11/06/p06008. URL: https://doi.org/10.1088%2F1748-0221%2F11%2F06%2Fp06008.

[60] G. Aad et al. *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System.* Tech. rep. CERN-LHCC-2013-018. ATLAS-TDR-023. Final version presented to December 2013 LHCC. 2013. URL: http://cds.cern.ch/record/1602235.

[61] G. Aad et al. “Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton proton collision data”. In: *The European Physical Journal C* 74.11 (2014). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3130-x. URL: http://dx.doi.org/10.1140/epjc/s10052-014-3130-x.

[62] "Muon reconstruction performance of the ATLAS detector in proton proton collision data at $\sqrt{s}$, volume=76, ISSN=1434-6052, url=http://dx.doi.org/10.1140/epjc/s10052-016-4120-y, DOI=10.1140/epjc/s10052-016-4120-y, number=5, journal=The European Physical Journal C, publisher=Springer Science and Business Media LLC, author=Aad, G. and Abbott, B. and Abdallah, J. and Abdinov, O. and Abeloos, B. and Aben, R. and Abolins, M. and AbouZeid, O. S. and Abraham, N. L. and et al., year=2016, month=May". In: ().

[63] P F Zema et al. *The GNAM monitoring system and the OHP histogram presenter for ATLAS*. Tech. rep. ATL-DAQ-CONF-2005-029. ATL-COM-DAQ-2005-031. CERN-ATL-COM-DAQ-2005-031. Geneva: CERN, 2005. URL: `https://cds.cern.ch/record/886320`.

[64] M. Aaboud et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton?proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 79.8 (2019). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-019-7140-6`. URL: `http://dx.doi.org/10.1140/epjc/s10052-019-7140-6`.

[65] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, 2008. URL: `https://cds.cern.ch/record/1099735`.

[66] T Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032014. DOI: `10.1088/1742-6596/119/3/032014`. URL: `https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032014`.

[67] T G Cornelissen et al. "The global χ2track fitter in ATLAS". In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032013. DOI: `10.1088/1742-6596/119/3/032013`. URL: `https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032013`.

[68] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, 2012. URL: `https://cds.cern.ch/record/1449796`.

[69] G. Aad et al. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *The European Physical Journal C* 74.10 (2014). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-3071-4`. URL: `http://dx.doi.org/10.1140/epjc/s10052-014-3071-4`.

[70] M. Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015?2016 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.03 (2019), P03017?P03017. ISSN: 1748-0221. DOI: `10.1088/1748-0221/14/03/p03017`. URL: `http://dx.doi.org/10.1088/1748-0221/14/03/P03017`.

[71] G. Aad et al. "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{\mathrm{s}} = 7$ TeV". In: *The European Physical Journal C* 73.3 (2013). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-013-2304-2`. URL: `http://dx.doi.org/10.1140/epjc/s10052-013-2304-2`.

[72] ATLAS Collaboration. *Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*. 2019. arXiv: `1910.04482 [hep-ex]`.

[73] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), pp. 063–063. DOI: `10.1088/1126-6708/2008/04/063`. URL: `https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063`.

[74] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *The European Physical Journal C* 72.3 (2012). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-012-1896-2`. URL: `http://dx.doi.org/10.1140/epjc/s10052-012-1896-2`.

[75] G. Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *The European Physical Journal C* 77.7 (2017). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-017-5004-5`. URL: `http://dx.doi.org/10.1140/epjc/s10052-017-5004-5`.

[76] *Impact of Alternative Inputs and Jet Grooming on Large-R Jet Performance*. Tech. rep. ATL-PHYS-PUB-2019-027. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2683619.

[77] G. Aad et al. "Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment". In: *Journal of High Energy Physics* 2012.5 (2012). ISSN: 1029-8479. DOI: 10.1007/jhep05(2012)128. URL: http://dx.doi.org/10.1007/JHEP05(2012)128.

[78] *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. Tech. rep. ATLAS-CONF-2016-035. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2200211.

[79] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The catchment area of jets". In: *Journal of High Energy Physics* 2008.04 (2008), pp. 005–005. DOI: 10.1088/1126-6708/2008/04/005. URL: https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F005.

[80] Andrew J. Larkoski, Ian Moult, and Duff Neill. "Power Counting to Better Jet Observables". In: *JHEP* 12 (2014), p. 009. DOI: 10.1007/JHEP12(2014)009. arXiv: 1409.6298 [hep-ph].

[81] Andrew J. Larkoski, Ian Moult, and Duff Neill. "Analytic Boosted Boson Discrimination". In: *JHEP* 05 (2016), p. 117. DOI: 10.1007/JHEP05(2016)117. arXiv: 1507.03018 [hep-ph].

[82] *Improving jet substructure performance in ATLAS using Track-CaloClusters*. Tech. rep. ATL-PHYS-PUB-2017-015. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2275636.

[83] Georges Aad et al. "Search for diboson resonances in hadronic final states in 139 fb^{-1} of *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 09 (2019), p. 091. DOI: 10.1007/JHEP09(2019)091. arXiv: 1906.08589 [hep-ex].

[84] "Performance of b-jet identification in the ATLAS experiment". In: *Journal of Instrumentation* 11.04 (2016), P04008. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/04/p04008. URL: http://dx.doi.org/10.1088/1748-0221/11/04/P04008.

[85] G. Aad et al. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$=13 TeV". In: *The European Physical Journal C* 79.11 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7450-8. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7450-8.

[86] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. ATL-PHYS-PUB-2017-013. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2273281.

[87] *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. ATL-PHYS-PUB-2017-011. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2270366.

[88] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. ATL-PHYS-PUB-2018-025. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2645405.

[89] Andreas Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*. 2007. arXiv: physics/0703039 [physics.data-an].

[90] *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. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2268678.

[91] ATLAS Collaboration. *Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector*. ATL-PHYS-PUB-2017-009. URL: https://cds.cern.ch/record/2263679.

[92] *Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector*. Tech. rep. ATL-PHYS-PUB-2017-017. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2275641.

[93] *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*. Tech. rep. ATL-PHYS-PUB-2015-027. Geneva: CERN, 2015. URL: `http://cds.cern.ch/record/2037904`.

[94] *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ TeV. Tech. rep. ATLAS-CONF-2017-029. Geneva: CERN, 2017. URL: `http://cds.cern.ch/record/2261772`.

[95] ATLAS Collaboration. *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. ATL-PHYS-PUB-2017-003. 2017. URL: `https://cds.cern.ch/record/2255226`.

[96] *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*. Tech. rep. ATL-PHYS-PUB-2019-033. Geneva: CERN, 2019. URL: `http://cds.cern.ch/record/2688062`.

[97] B. Mehlig. *Artificial Neural Networks*. 2019. arXiv: `1901.05639 [cs.LG]`.

[98] Geoffrey Hinton Yann LeCun Yoshua Bengio. "Deep learning". In: (2015). URL: `https://www.nature.com/articles/nature14539`.

[99] Zachary C. Lipton, John Berkowitz, and Charles Elkan. *A Critical Review of Recurrent Neural Networks for Sequence Learning*. 2015. arXiv: `1506.00019 [cs.LG]`.

[100] Sepp Hochreiter and Jürgen Schmidhuber. "Long short-term memory". In: *Neural computation* 9.8 (1997), pp. 1735–1780.

[101] Zachary Chase Lipton. "A Critical Review of Recurrent Neural Networks for Sequence Learning". In: *CoRR* abs/1506.00019 (2015). arXiv: `1506.00019`. URL: `http://arxiv.org/abs/1506.00019`.

[102] Yonghui Wu et al. "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation". In: *CoRR* abs/1609.08144 (2016). arXiv: `1609.08144`. URL: `http://arxiv.org/abs/1609.08144`.

[103] François Chollet et al. *Keras*. `https://keras.io`. 2015.

[104] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN: 1441412697.

[105] URL: `https://keras.io/layers/core/`.

[106] URL: `https://keras.io/optimizers/`.

[107] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15 (2014), pp. 1929–1958. URL: `http://jmlr.org/papers/v15/srivastava14a.html`.

[108] Georges Aad et al. "Search for the electroweak diboson production in association with a high-mass dijet system in semileptonic final states in *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D100.3 (2019), p. 032007. DOI: `10.1103/PhysRevD.100.032007`. arXiv: `1905.07714 [hep-ex]`.

[109] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. ATLAS-CONF-2019-021. 2019. URL: `https://cds.cern.ch/record/2677054`.

[110] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *Journal of Instrumentation* 13.07 (2018), P07017–P07017. DOI: `10.1088/1748-0221/13/07/p07017`. URL: `https://doi.org/10.1088%2F1748-0221%2F13%2F07%2Fp07017`.

[111] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: `10.1016/S0168-9002(01)00089-4`.

[112] Torbjorn Sjostrand. "PYTHIA 8 Status Report". In: (2008), pp. 726–732. arXiv: `0809.0303 [hep-ph]`.

[113] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[114] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[115] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 0902 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[116] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 1006 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[117] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[118] Pierre Artoisenet et al. "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations". In: *JHEP* 1303 (2013), p. 015. DOI: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph].

[119] *ATLAS Run 1 Pythia8 tunes*. Tech. rep. ATL-PHYS-PUB-2014-021. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1966419.

[120] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: https://cds.cern.ch/record/2216168.

[121] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[122] ATLAS Collaboration. *Performance of electron and photon triggers in ATLAS during LHC Run 2*. 2019. arXiv: 1909.00761 [hep-ex].

[123] M. Aaboud et al. "Measurement of the cross-section for electroweak production of dijets in association with a Z boson in pp collisions ats=13 TeVwith the ATLAS detector". In: *Physics Letters B* 775 (2017), 206?228. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2017.10.040. URL: http://dx.doi.org/10.1016/j.physletb.2017.10.040.

[124] M. Aaboud et al. "Measurement of the Higgs boson coupling properties in the $H \to ZZ \to 4l$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2018.3 (2018). ISSN: 1029-8479. DOI: 10.1007/jhep03(2018)095. URL: http://dx.doi.org/10.1007/JHEP03(2018)095.

[125] *ELI5*. https://eli5.readthedocs.io/en/latest/,

[126] ATLAS Collaboration. "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} =$ 13 TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 117.18 (2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. arXiv: 1606.02625 [hep-ex].

[127] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[128] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[129] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton proton collision data at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[130] ATLAS collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data". In: (2019). arXiv: 1908.00005 [hep-ex].

[131] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[132] ATLAS Collaboration. "In situ calibration of large-R jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector". In: *Eur. Phys. J.* C 79 (2019), p. 135.

[133] ATLAS Collaboration. *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*. ATLAS-CONF-2018-001. 2018. URL: https://cds.cern.ch/record/2306649.

[134] ATLAS Collaboration. "Measurements of *b*-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s}$ =13 TeV". In: *JHEP08* 089 (2018). arXiv: 1805.01845 [hep-ex].

[135] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV". In: *Eur. Phys. J.* C 78.903 (2018). arXiv: 1802.08168 [hep-ex].

[136] John M. Campbell, R. Keith Ellis, and Ciaran Williams. "Vector boson pair production at the LHC". In: *JHEP* 07 (2011), p. 018. DOI: 10.1007/JHEP07(2011)018. arXiv: 1105.0020 [hep-ph].

[137] S. Catani et al. "QCD matrix elements + parton showers". In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231 [hep-ph].

[138] Stefan Hoeche et al. "QCD matrix elements and truncated showers". In: *JHEP* 05 (2009), p. 053. DOI: 10.1088/1126-6708/2009/05/053. arXiv: 0903.1219 [hep-ph].

[139] M. Bahr et al. "Herwig++ physics and manual". In: *Eur. Phys. J. C* 58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[140] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[141] L.A. Harland-Lang et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". In: *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].

[142] V. Barger et al. "Strong W^+W^+ scattering signals at pp supercolliders". In: *Phys. Rev. D* 42 (9 1990), pp. 3052–3077. DOI: 10.1103/PhysRevD.42.3052. URL: https://link.aps.org/doi/10.1103/PhysRevD.42.3052.

[143] ATLAS Collaboration. *Proposal for particle-level object and observable definitions for use in physics measurements at the LHC*. ATL-PHYS-PUB-2015-013. 2015. URL: https://cds.cern.ch/record/2022743.

[144] G. Aad et al. "Combined search for the Standard Model Higgs boson inppcollisions ats=7??TeVwith the ATLAS detector". In: *Physical Review D* 86.3 (2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.032003. URL: http://dx.doi.org/10.1103/PhysRevD.86.032003.

[145] Lorenzo Moneta et al. *The RooStats Project*. 2010. arXiv: 1009.1003 [physics.data-an].

[146] Wouter Verkerke and David P. Kirkby. "The RooFit toolkit for data modeling". In: (2003). arXiv: physics/0306116 [physics.data-an].

[147] Alexander L. Read. "Presentation of search results: The CL(s) technique". In: *J. Phys.* G28 (2002). [,11(2002)], pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[148] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[149] ATLAS Collaboration. "Combined search for the Standard Model Higgs boson in *pp* collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector". In: *Phys. Rev. D* 86 (2012), p. 032003. DOI: 10.1103/PhysRevD.86.032003. arXiv: 1207.0319 [hep-ex].

[150] Lorenzo Moneta et al. "The RooStats Project". In: (2010). arXiv: 1009.1003 [physics.data-an].

[151] ATLAS Collaboration. *Procedure for the LHC Higgs boson search combination in summer 2011*. ATL-PHYS-PUB-2011-011. 2011. URL: https://cds.cern.ch/record/1375842.

Acknowledgements

As each beautiful story also this thesis needed an end... I would like to write still few words to try to thanks all people that have a piece of this work.

Innanzitutto, vorrei ringraziare il prof. Merola, mio relatore fin dalla tesi triennale, senza il quale il mio intero percorso nella Fisica delle Particelle non sarebbe mai iniziato; sono partito con lui dietro i banchi del suo corso di Elettromagnetismo ed Ottica per arrivare ad essere un dottorando delle esperimento ATLAS e collaborare con lui al corso di Fisica delle Particelle Elementari.

Vorrei poi ringraziare il Dott. Gianpaolo Carlino e l'intero gruppo ATLAS Napoli, che oltre ad avermi accettato in questo gruppo mi hanno sostenuto per lasciarmi inserire in una comunità ben più ampia, quella del CERN, potendo spendere più di anno in quel mondo fantastico che si trova sul confine franco/svizzero in corrispondenza di un anello di 27 km di circonferenza…

Non basterebbe un'altra intera tesi di ringraziamenti per i miei supervisors, il Dott. Francesco Conventi e la Dott.ssa Elvira Rossi, i quali mi hanno condotto ed aperto le porte di questo universo che è la ricerca in fisica delle alte energie. E' solo grazie a loro che, dopo 3-4 anni, sono nuovamente a riempire una nuova pagina di ringraziamenti. Come immaginavo al termine della laurea magistrale mi trovavo solo sulla linea di partenza di un percorso che mi avrebbe portato ad affrontare sfide, a "sbattere" la testa, così come solo nella ricerca può accadere, e, soprattutto, a crescere per comprendere il significato di essere in prima linea nella ricerca fondamentale, seppure sia ancora alla prime armi. Solo attraverso fatica, stanchezza, serate (e nottate) di lavoro, rinunce ma anche confronti, risate e cene insieme sono arrivato soddisfatto al termine di questa tesi, orgoglioso del lavoro svolto con loro e delle soddisfazioni raggiunte insieme che riescono sempre a stimolare a fare meglio e di più. Ho sempre cercato e cercherò sempre di imparare "u mestier" che, come credo sia giusto in questo mondo, non si finirà mai di imparare.

Un grande ringraziamento va al Dott. Attilio Picazio, che mi ha conosciuto come uno spaventato dottorando al primo anno, alla vigilia della prima presentazione nazionale, e con il quale ho avuto il piacere di condividere gioie e dolori del 2019 trascorso al CERN. Mi ha insegnato che, nonostante tutto, in un mondo fatto di numeri, codici e plots, siamo sempre persone, solo giovani ricercatori a volte, e che il rapporto umano deve sempre giocare il primo ruolo anche nelle collaborazioni scientifiche e competitive come quelle del CERN. Sono stati indispensabili consigli, rimproveri e anche birre ad r1 dopo ore di meeting con le mie immancabili 40 slides di giovedì alle 17 del pomeriggio.

Un grazie anche ai miei analysis contacts del caro "DBL VV semi-leptonic analysis", Lailin & Takuya, che nonostante tensioni e fatiche scientifiche hanno giocato un ruolo nella mia crescita di dottorando (thanks guys!).

Ringraziamento anche ai compagni di laboratorio, Francesco, Marco e Silvia, con i quali in modi e tempi diversi abbiamo condiviso tempo, caffè e pranzi sulla tastiera; un ulteriore grazie a Francesco, che costituisce sempre la mia voce della coscienza in quel di msa e che ha la pazienza di sopportare i miei sfoghi e che mi stimola in confronti di ogni genere su fisica e non solo.

Ringraziamento caloroso ai miei amici di sempre, Pietro, Federica e Anna Rita, con i quali ho condiviso e condivido tanti "scampoli di esistenza" ormai da quasi 15 anni, anche se ormai lontani, restano fondamentali i nostri momenti di quotidiana condivisione e di incontri e cene quando possibile.

Un enorme e sincero grazie a mia madre, che deve sopportare tutti i miei difetti, sfoghi, nervosismi e ormai anche dimenticanze dovute alla mia mente sempre piu spesso trasportata altrove, nel mio lavoro e passione. Un ringraziamento finale ad una persona senza la quale non avrei mai potuto immaginare di ritrovarmi a questo punto della mia vita con un enorme bagaglio di esperienze e passioni e di poter vivere questo sogno, lui sa chi è.