

Università degli Studi di Napoli “Federico II”

Scuola Politecnica e delle Scienze di Base
Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”



Laurea Magistrale in Fisica

**Search for electroweak diboson production
in $ZV \rightarrow llqq$ ($V=Z/W$) final states with the
ATLAS detector at the LHC**

*Ricerca di produzione elettrodebole di sistemi con due bosoni nei canali
 $ZV \rightarrow llqq$ nell'esperimento ATLAS a LHC*

Relatori:

Prof. Leonardo Merola
Dott.ssa Elvira Rossi

Candidato:

Sara Pinto
Matricola N94000377

A.A. 2019/2020

*The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say.*

J.R.R. Tolkien, *The Fellowship of the Ring*, Book I, Chapter 1

Contents

Introduction	i
1 The Standard Model of the elementary particles	1
1.1 Elementary particles	1
1.1.1 Matter constituents	2
1.1.2 Leptons and quarks	2
1.1.3 Force-carriers	3
1.2 The Quantum Field Theory and the Yang-Mills theory basis	4
1.2.1 Quantisation of classical fields	4
1.2.2 Global and local invariance of the Lagrangian	5
1.2.3 Local invariance of the Lagrangian under $SU(N)$	7
1.3 The Standard Model	9
1.3.1 The Electroweak theory	9
1.3.2 The Brout–Englert–Higgs mechanism	11
1.3.3 The Quantum Chromodynamics	13
1.3.4 Fermions masses	15
1.3.5 Vector Boson Scattering	16
2 ATLAS Experiment at the LHC, CERN	18
2.1 The Large Hadron Collider	18
2.1.1 Run 1	19
2.1.2 Run 2	20
2.1.3 Run 3 and High Luminosity LHC	21
2.2 ATLAS: A Toroidal Lhc ApparatuS	23
2.2.1 ATLAS Coordinate System	24
2.2.2 ATLAS magnets	26
2.2.3 The Inner Detector	26
2.2.4 The Calorimetric System	28
2.2.4.1 Electromagnetic Calorimeter	29
2.2.4.2 Hadronic Calorimeters	30
2.2.5 The Muon Spectrometer	30
2.3 The ATLAS Trigger System	32
2.3.1 The Electron Trigger	34
2.3.2 The Muon Trigger	35
3 Physical Object Reconstruction	37
3.1 Electrons reconstruction	37
3.1.1 Electron identification	39
3.2 Muons reconstruction	40
3.2.1 Muon identification	41
3.3 Jets reconstruction	44
3.3.1 Small- R jets	46
3.3.1.1 Reconstruction	47

3.3.1.2	Jet energy scale calibration	47
3.3.1.3	Jet energy resolution	47
3.3.2	Large- R jets	48
3.3.2.1	Vector boson tagging	51
4	Search for electroweak diboson production in semileptonic final states	53
4.1	Data and Monte Carlo samples	54
4.1.1	Data samples	54
4.1.2	Monte Carlo simulation of signal and background processes	54
4.2	Object definition	57
4.2.1	Leptons	57
4.2.2	Small- R jets	58
4.2.3	Large- R jets	58
4.2.4	Overlap Removal	58
4.3	The cut-based approach	59
4.3.1	$Z \rightarrow ll$ search	59
4.3.2	VBS Topology: tagging jets	59
4.3.3	Selection of $V_{\text{had}} \rightarrow J$ candidate (merged selection)	60
4.3.4	Selection of $V_{\text{had}} \rightarrow jj$ candidate (resolved selection)	60
4.3.5	Control regions	60
4.4	The Recurrent Neural Network approach	62
4.4.1	RNN architecture and training	62
4.4.2	VBS topology categorization	65
4.4.2.1	Preliminary study of signal leading jet p_T	67
4.4.2.2	Signal efficiency and significance with respect to cut-based approach	67
4.4.2.3	The signal jets pairing algorithm in the resolved regime	68
4.4.2.4	RNN score tight cuts	72
4.4.3	Training in signal regions	79
4.4.3.1	RNN performance at high m_{VV}	88
4.4.3.2	Signal jets pairing: an alternative algorithm	100
5	Statistical analysis	111
5.1	Previous run statistical approach	111
5.1.1	Likelihood function and statistical hypothesis test	111
5.1.2	Fit inputs	112
5.1.3	Statistical test	113
5.2	Current approach	114
5.2.1	Expected significance	114
5.2.2	aQCGs search strategy	115
	Conclusions	115
	Bibliography	117

Introduction

My Thesis work is part of a search for electroweak diboson production in semileptonic final states currently performed by the ATLAS Collaboration. The data exploited in the search, corresponding to an integrated luminosity of 139 fb^{-1} , were recorded with the ATLAS detector from the Large Hadron Collider proton-proton collisions that occurred at a centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ from 2015 to 2018.

The work here presented focuses on a specific diboson process, namely the scattering of vector bosons. Vector boson scattering searches play a twofold role in the modern particle physics scenario. On one hand, they represent a key tool for probing the Standard Model of the weak, strong and electromagnetic interactions. In fact, as opposed to the electromagnetic force carriers, namely the photons, the peculiar non-Abelian gauge structure of the Standard Model allows for vector bosons V (the weak force carriers, $V = Z/W$) auto-interactions; moreover, for such processes to avoid unitarity violation, the Higgs boson is needed to participate in the interaction.

On the other hand, diboson production processes can be exploited to search for new physics. Indeed, the Standard Model has successfully predicted a wide range of experimental results, the discovery of the Higgs boson in 2012 being one such example, and it has provided the scientific community with a thorough comprehension of what matter is made of and how subatomic particles interact. Nonetheless, there are still many unanswered questions: why are there three generations of leptons and quarks? Why are their masses so different from each other? Where does the unbalance between matter and anti-matter come from? How can gravitation be included in a unified theory? What is this dark energy which accounts for so much of the universe we cannot see?

One widely held opinion is that the Standard Model is nothing but a low-energy manifestation of a more general theory. Vector boson scattering is involved with gauge couplings that are predicted to be modified with respect to the Standard Model prediction at energies of order 1 TeV. The remarkable technological progress made in the last decades opens the possibility for the largest accelerator ever built, the Large Hadron Collider, to observe eventual anomalous couplings or, otherwise, to set limits of unprecedented precision on their existence.

The search for electroweak diboson production which I took part in is currently at a very preliminary stage. It aims to investigate the VV semileptonic decay channels, i.e. final states in which one boson decays hadronically, while the other boson decays leptonically. My contribution regarded the optimisation of the event selection for the 2-lepton channel, with two charged leptons in the final state ($ZV \rightarrow llqq$), by the employment of a Recurrent Neural Network, with plans to considerably enhance the expected signal efficiency and the expected significance.

The present Thesis consists of five chapters. A summary of the Standard Model theoretical aspects, including an overview of vector boson scattering, is given in the first chapter. The second and third chapters deal with the description of the ATLAS detector, with all its sub-constituents, alongside with the main techniques employed to reconstruct physical objects from the detector signals. The fourth chapter gives an insight into the analysis, with particular emphasis on my contribution. In the final chapter, a summation of the statistical methods which can be exploited to extract the cross-section measurement as well as to set limits on anomalous quartic gauge couplings is provided.

Chapter 1

The Standard Model of the elementary particles

As of today, Particle Physics embodies man's most elegant and efficient effort made in order to address the question of what the universe is formed from. The idea of matter being indefinitely divisible, although beautifully representative of some simplicity behind the complexity of nature, persisted until the discovery of the atom, which paved the way for the concept of *elementary particles* as being the apparently structure-less fundamental units forming the universe.

Investigating the nature of such particles does not merely imply some sort of basic classification; on the contrary, it is inevitably tied to the study of particles mutual interactions, occurring via the *fundamental forces*. Consistently with nature's predilection for economy, there appears to exist only four of such forces: gravitation, whose theoretical formulation was given by Isaac Newton and Albert Einstein; electromagnetism, described by James Clerk Maxwell's and Richard Feynman's theories; the strong and the weak interactions, firstly described by Hideki Yukawa and Enrico Fermi's models, respectively. While both gravitational and electromagnetic force ranges of action are large enough to let us have direct experience of their effects in our everyday lives, the strong and the weak forces exert their influence over very short distances (from 10^{-16} cm to 10^{-13} cm) and are therefore mainly involved in nuclear and sub-nuclear phenomenology.

Pursuing the idea of a unified theory, physicists have made remarkable progress in finding similarities between three of the fundamental forces, so that nowadays the electromagnetic, the strong and the weak interactions can be described through the same mathematical framework, that is *Quantum Field Theory* (QFT). More specifically, the form and the properties of interactions arise from requiring as valid the *principle of local gauge invariance*, which means that the equations describing particle physics phenomena keep the same form after applying some transformation to the particle *fields*. The Standard Model (of the strong, weak and electromagnetic interactions) is the simplest realization of such principle.

Despite actually not providing a full unification to the three fundamental interactions it is involved with (not to mention the fact that gravitation is completely excluded from the picture), the Standard Model has successfully predicted the existence of many particles, including the massive gauge bosons, Z and W , alongside with the Higgs boson, whose observation has confirmed the Standard Model to be a self-consistent theory.

In this chapter a summary of the main theoretical aspects of the Standard Model is displayed.

1.1 Elementary particles

Each particle is described by a set of quantum numbers, often referred to as *internal*, since they do not deal with the kinematic state of the particle itself. The whole set of elementary particles can be split into two fundamental categories, according to a quantum number called *spin* (whose value shall be from now on expressed in units of Planck's constant $\hbar$):

- fermions: particles whose spin is a semi-integer number and which obey the Fermi-Dirac statistics;
- bosons: particles whose spin is an integer number and which obey the Bose-Einstein statistics.

As a general rule, to every particle corresponds a so-called antiparticle, a copy which differs only in the sign of its quantum numbers (first of all, the electric charge). In some cases, the particle cannot be distinguished from its antiparticle.

1.1.1 Matter constituents

Matter constituents happen to be a few: the atom is the result of negatively charged electrons (e^-) being bound to positively charged nuclei by the electromagnetic force (a low-energy manifestation of QED, short for *Quantum Electrodynamics*). Both electrons and nuclei are capable of interacting electromagnetically as a result of their electric charge (-1 for the electron and $+1$ for the proton in units of the elementary charge $e \approx 1.6 \times 10^{-19}$ C).

Electrons are fermions with $s = 1/2$ and belong to the lepton class. The nuclei themselves are nothing different from a bunch of composite particles, protons and neutrons, held together by the nuclear strong force, the latter being a large-scale manifestation of *Quantum Chromodynamics* (QCD). Looking at even smaller scales, protons and neutrons are found to be formed from half-spin fermions called quarks (of type up and down, respectively denoted by u and d).

The notion of what an elementary particle is changes in step with the smallest distance we are able to access, or equivalently with the rise in power of the sources deployed to investigate the internal structure of matter. More precisely, the higher the energy of the probe particle is, the smaller is the wavelength associated to the latter and, consequently, the resolution obtained in the measure. At the smallest distances currently probed by the highest-energy accelerators, leptons and quarks seem not to exhibit further internal structure: as far as we know, the world around us is entirely made of half-spin fermions, namely leptons and quarks, interacting via the electromagnetic and the strong forces (Figure 1.1).

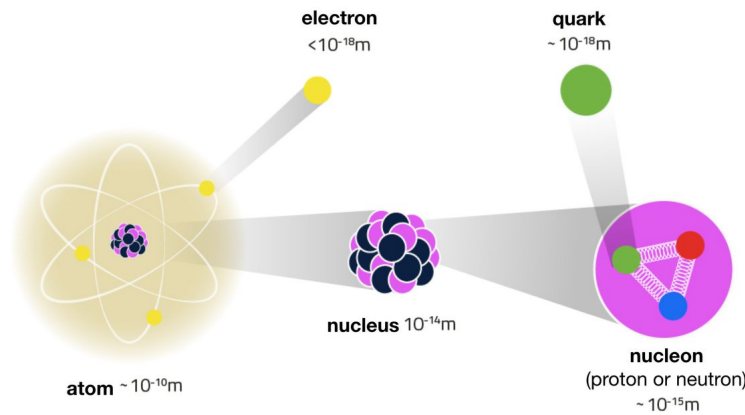


Figure 1.1: Schematic portrait of the atom.

1.1.2 Leptons and quarks

In contrast to the strong force, the weak force (as its name suggests) is not intense enough to hold different particles together to produce bound states, i.e. matter particles; it is mainly involved in radioactive decays and nuclear fusion. Nevertheless, the weak force plays a crucial role in the Standard Model, appearing to be partially unified to the electromagnetic one in what is called *electroweak force*. In processes involving the weak force, another lepton is often produced: the neutrino (ν).

The electron, together with the (electronic) neutrino (ν_e), belongs to a so-called lepton doublet. Leptons are half-spin fermions interacting electromagnetically and weakly, but never strongly. There exist three families, or *generations*, of lepton doublets, the first being composed of the electron and its neutrino, while the muon (μ^-) and the tau (τ^-) leptons, together with their own neutrinos (ν_μ and ν_τ), belong to the second

and third generation, respectively. The masses of the charged leptons (whose electric charge is -1) increase from the first to the last generation, whereas all neutrinos are nearly mass-less. Neutrinos are also electrically uncharged, therefore they can only interact via the weak force.

Leptons and anti-leptons are respectively assigned a lepton number $L = 1$ and $L = -1$. In order not to be forbidden, a physical reaction must always ensure that the total lepton number equals zero. Each lepton doublet is also given a so-called *family* lepton number $L_l = 1$, with $l = e, \mu, \tau$ ($L_l = -1$ for anti-leptons), which empirically appears to be always conserved in physical processes: specifically, the number of leptons minus the number of anti-leptons within the same generation does not change from the initial to the final state of any reaction.

Lepton doublets participate in the arrangement of *fermions generations* together with quarks. Similarly to leptons, quarks are half-spin fermions. They can interact electromagnetically (being provided with fractional electric charges), strongly and weakly. As well as leptons, they are organized into doublets, each doublet containing a particle with an electric charge of $+2/3$ and a particle with an electric charge of $-1/3$: we have the up-quark and the down-quark (u and d) for the first generation, the charm-quark and the strange-quark (c and s) for the second generation, and the top-quark and the bottom- (or beauty-) quark (t and b) for the third generation: a schematic illustration is given in Figure 1.2 (including the antiparticles as well).

Quarks are given a baryon number $B = 1/3$ ($B = -1/3$ for anti-quarks) to account for the experimental fact that they are only found within composite *hadrons*, the latter being formed from either three quarks states (*barions*, identified by $B = 1$ and an integer electric charge) or quark-anti-quark (not necessarily of the same type) states (*mesons*, identified by $B = 0$ and an integer electric charge).

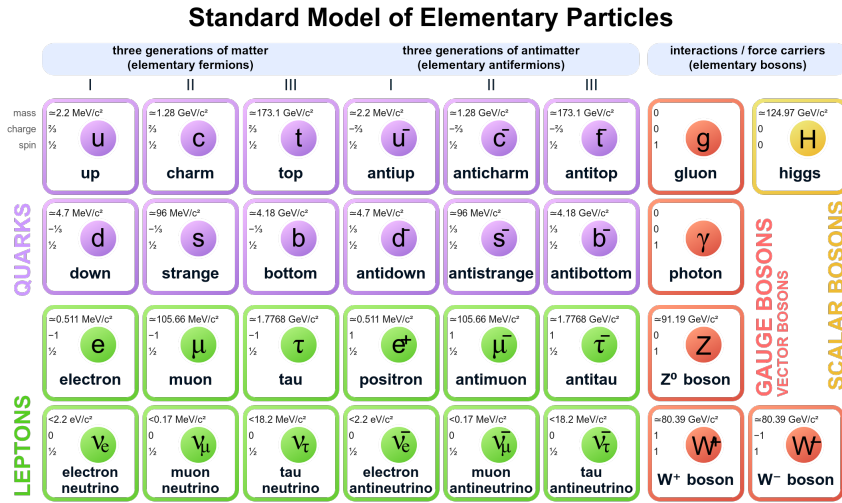


Figure 1.2: The elementary particles and antiparticles, divided into quarks, leptons and force-carrying particles, together with the scalar Higgs boson.

1.1.3 Force-carriers

In classical physics, matter and forces have always belonged to two completely separated worlds. Despite some initial difficulties in digesting the idea that interactions at-a-distance would be possible by means of *fields* permeating the space surrounding massive (in the case of gravitation) and electronic and magnetic (in the case of electromagnetism) bodies, the distinction between matter and forces seemed to be clear until the early twentieth century, when quantum mechanics made its appearance in the modern physics scenario. On one hand, experiments like the photoelectric effect and the Compton effect unmistakably proved light to have a particle-like nature (i.e. the photon) beside its wave-like one. On the other hand, the interference and the diffraction observed for electrons made it clear that particle-like objects displayed a wave-like behaviour as well.

At present, interactions between particles are not described by classical fields any more: *quantum* fields are indeed responsible for interactions at-a-distance. Let us consider a particle receiving a push in some direction: to produce an effect on a particle nearby, the first particle emits a quantum which is in charge of carrying the amount of energy and momentum to be transferred to the second particle, so that the two particles are actually capable of interacting at-a-distance by means of an intermediate: the force carrier.

But how come a particle can emit an energy quantum of mass m , or equivalently lose an energy amount of $\Delta E = mc^2$, without energy conservation being violated? Actually, according to Heisenberg's principle, energy conservation violation is allowed as long as it occurs in a fraction of time satisfying the condition $\Delta t \Delta E \geq \hbar/2$. Since they cannot directly being experimentally observed, force carriers are referred to as *virtual* quanta.

An example of interaction by exchange of virtual quanta is schematically represented in Figure 1.3 by a Feynman diagram, where two electrons (the upper and the lower straight lines) interact by exchanging a photon (the middle wavy line). Considering time as running from left to right, two time-ordered diagrams are equally probable, as the exchanged particle is not observed: therefore, the combination of the two virtual processes is what is physically meaningful.

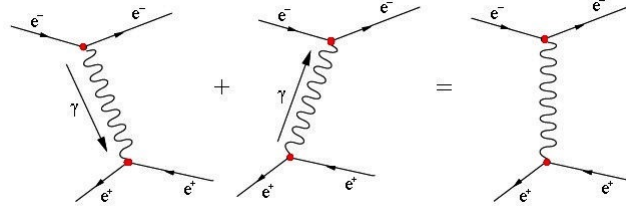


Figure 1.3: The electromagnetic interaction between two electrons by exchange of a photon.

The existence of virtual particles at each point of space and time is encoded in the concept of field, whose quantized fluctuations with respect to its ground state result in having particle-like features.

The force carriers are summarized in Figure 1.2, alongside with their basic properties. They are all bosons with $s = 1$. As we shall see later on, there exist eight different gluons (g), which are carriers of the strong force and exhibit zero mass, as well as the photon (γ), which is responsible for the electromagnetic force. In contrast, the three carriers of the weak force (Z^0 , W^+ , W^-) have non-vanishing values for masses and electric charges. The gauge boson mass is crucially related to the force range of action through the condition $c\Delta t = \hbar/mc$, that is the maximum distance the boson can travel within Δt accordingly to Heisenberg's principle.

1.2 The Quantum Field Theory and the Yang-Mills theory basis

As we mentioned above, the Standard Model lies on Quantum Field Theory (QFT) principles and mathematical framework. As such, it is involved with quantum fields which arise from a procedure called *quantisation* of classical fields.

1.2.1 Quantisation of classical fields

Gravitation and classical electrodynamics are both described by a non-quantum formalism, in which classical fields play a crucial role in describing physical quantities: at each point of space and time, a physical quantity, that being of scalar, vectorial or tensor nature, is assigned.

For quantum phenomena to be taken into account, a new formalism is needed: to describe the position and the momentum of a particle (at each point in time), the corresponding classical variables (fields) are substituted by mathematical operators, which exert their action over a Hilbert space $\mathcal{H}$. This substitution is called *first quantisation*. At this point, the possible states of a physical system are represented by vectors (usually denoted by $\psi(x)$) in the Hilbert space, and the probability density for such system to be in one of these states equals the norm square of the corresponding vector, $|\psi(x)|^2$.

Unfortunately, when it comes to relativity, quantum mechanics alone fails to provide a satisfactory framework. In order to conciliate special relativity and quantum mechanics, a *second quantisation* must occur: the probability amplitude $\psi(x)$ is substituted by a field operator $\hat{\phi}(x)$, which now exerts its action over a Fock space $\mathcal{F}$, i.e., a space containing all the possible multi-particle states of a physical system. Such states arise from the subsequent action of creation ($\hat{a}$) or annihilation ($\hat{a}^\dagger$) operators over the quantum vacuum state of the system $|0\rangle$, the latter corresponding to a minimum energy configuration rather than to an absence of particles.

1.2.2 Global and local invariance of the Lagrangian

A field operator $\phi(x)$ ¹ is nothing but a linear combination of creation and annihilation operators. The Lagrangian L is a function of the fields $\phi_j(x)$ (more than one field is often required, depending on the degrees of freedom of the system to be described) which summarizes the dynamics of the system. L is usually written in terms of a Lagrangian density $\mathcal{L}$, which itself depends on the fields $\phi_j(x)$ and their gradients $\frac{\partial \phi_j}{\partial x^\mu} \equiv \partial_\mu \phi_j(x)$:

$$L(t) = \int \mathcal{L}(\phi_j(x), \partial_\mu \phi_j(x)) d^3 \mathbf{x}. \quad (1.1)$$

We denote by x the space-time coordinates $(ct, x, y, z) = (ct, \mathbf{x})$ and use the natural units system, where $\hbar = 1$ and $c = 1$. The Lagrangian density $\mathcal{L}$ obeys the Euler-Lagrange equations (one for each field $\phi_j(x)$):

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

The sum over repeated indices is implied. These equations descend from a variational principle, i.e., from requiring that the solutions $\phi_j(x)$ are such that the action $\mathcal{S}$, defined as:

$$\mathcal{S} \equiv \int_{t_a}^{t_b} L(t) dt = \int_{x_a}^{x_b} \mathcal{L}(\phi_j(x), \partial_\mu \phi_j(x)) d^4 x \quad (1.3)$$

is stationary with respect to arbitrary variations of the solutions themselves, provided that those variations vanish at x_a and x_b . While the dynamic equations arise from a fundamental principle, the form of $\mathcal{L}$ is only constrained to satisfy some basic conditions in order to be physically meaningful, such as being Lorentz-invariant. This invariance is correlated to space-time coordinates transformations, but there exist other types of transformations, such as *gauge* ones, which act on the fields.

Let us consider the Lagrangian density $\mathcal{L}_0(x)$ for a field $\psi(x)$:

$$\mathcal{L}_0(x) = i\bar{\psi}(x)\gamma^\mu \partial_\mu \psi(x) - m\bar{\psi}(x)\psi(x) \quad (1.4)$$

where $\bar{\psi} \equiv \psi^\dagger \gamma^0$ and $\gamma^\mu = \{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$ are the gamma matrices satisfying the anti-commutation rules $\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}$, $g^{\mu\nu}$ being the Minkowski space-time metric tensor. Since the Euler-Lagrange equation for this field is the well-known Dirac equation:

$$(i\not{\partial} - m)\psi = 0 \quad (1.5)$$

where $\not{\partial} \equiv \gamma^\mu \partial_\mu$, we know that $\mathcal{L}_0$ is suited to describing half-spin fermions of mass m . Let us consider the gauge transformation:

$$\psi \rightarrow \psi' = \psi e^{-iq\theta}, \quad (1.6)$$

$$\bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} e^{iq\theta}. \quad (1.7)$$

Since the parameter defining the above written transformation, namely θ , does not depend on time-space coordinates, the transformation itself is called *global*.

The set of all possible transformations of the kind 1.6 (as θ changes) forms an abelian Lie group² called U(1): it consists of *unitary* transformations and it is mono-dimensional. It is a simple exercise to

¹For simplicity, the hat symbol used to denote operators is omitted.

²A group G is a set equipped with a binary operation satisfying four conditions:

verify that $\mathcal{L}_0$ is invariant under U(1) gauge transformations, i.e. $\mathcal{L}'_0 = \mathcal{L}_0$. As a consequence of this global symmetry, Noether's theorem applies to $\mathcal{L}_0$, ensuring the existence of a current $j^\mu = q\bar{\psi}\gamma^\mu\psi$ which satisfies the conservation law:

$$\partial_\mu j^\mu = 0. \quad (1.8)$$

The current j^μ is nothing but a relativistic version (for half-spin particles) of the electromagnetic current density $\mathbf{j} = \rho \mathbf{v}$. In this case, $j^0 = q\bar{\psi}\gamma^0\psi = q\psi^\dagger\gamma^0\psi = q\psi^\dagger\psi = q|\psi|^2$ can be identified with the spatial charge density³, so that:

$$q = \int_{\mathbb{R}^3} j^0 d^3\mathbf{x}. \quad (1.9)$$

From (1.8), one can write $\partial_0 j^0 = -\partial_i j^i$ and integrate over $\mathbb{R}^3$. Provided that the field $\psi(x)$ vanishes for $\mathbf{x} \rightarrow \infty$ quickly enough, the conservation law results in the conservation of the total electric charge, as expected:

$$\frac{dq}{dt} = - \int_{\mathbb{R}^3} \nabla \cdot \mathbf{j} d^3\mathbf{x} = - \int_S \mathbf{j} \cdot \mathbf{n} dS = 0 \Rightarrow q = \text{constant}. \quad (1.10)$$

In summation, the *global* invariance of $\mathcal{L}_0$ under U(1) leads to the existence of a conserved current j^μ as a consequence of Noether's theorem. The conservation law 1.8 implies the conservation of the electric charge.

If the parameter θ defining U(1) changes with x , the transformations of U(1) are called *local* and $\mathcal{L}_0$ is no longer invariant under U(1): $\mathcal{L}'_0 = \mathcal{L}_0 + j^\mu \partial_\mu \theta$. Since $\mathcal{L}_0$ is a scalar quantity, a term $\mathcal{L}_{\text{int}}$ can be added to ensure that $\mathcal{L} \equiv \mathcal{L}_0 + \mathcal{L}_{\text{int}}$ is locally invariant under U(1). For such invariance to be preserved, one can see that

$$\mathcal{L}_{\text{int}} = -j^\mu A_\mu \quad (1.11)$$

with $A^\mu(x)$ being a quadri-vector field such that, when a local transformation $e^{-iq\theta(x)}$ of U(1) is performed on $\psi(x)$, it undergoes the transformation:

$$A_\mu' = A_\mu + \partial_\mu \theta. \quad (1.12)$$

With these assumptions, $\mathcal{L} = \mathcal{L}_0 - j^\mu A_\mu$ satisfies the condition $\mathcal{L}' = \mathcal{L}$. It is of the utmost importance to notice that a massive term for the *gauge field* A_μ would break the gauge invariance for which A_μ was introduced in the first place. Furthermore, it can be shown that a quadri-vector field is nothing but a $s = 1$ representation of the Lorentz group $SO(1, 3)_c$ ⁴. In other words, the particle associated with this new field is a *mass-less gauge boson*.

-
- closure: performing the operation on members of the set produces a member of the same set;
 - associativity: denoting the binary operation by $\star$ and three arbitrary elements of the group by g_1, g_2, g_3 , it is true that $(g_1 \star g_2) \star g_3 = g_1 \star (g_2 \star g_3)$;
 - identity: there exists the identity element, denoted by 1, such that $1 \star g = g \star 1 = g, \forall g \in G$;
 - invertibility: $\forall g \in G, \exists! g^{-1} : g \star g^{-1} = g^{-1} \star g = 1$.

G is abelian if $g_1 \star g_2 = g_2 \star g_1 \forall g_1, g_2 \in G$. A continuous group is such that for each and every one of its elements, there exists only one counterpart point on a differential variety. If, defining a coordinate system, each element of the group can be identified by a n -tuple $(a_1, \dots, a_n) = \mathbf{a} \in \mathbb{R}^n$, G is said to have dimension n . If $g_1 \rightarrow \mathbf{a}, g_2 \rightarrow \mathbf{b}$, then $g_3 = g_1 \star g_2 \rightarrow \mathbf{c} = \mathbf{f}(\mathbf{a}, \mathbf{b})$, with $\mathbf{f} = (f_1, \dots, f_n)$. If f_i is an analytic function $\forall i$, G is called a Lie group. To put it simply, a Lie group is such that each of its elements can be described by a n -tuple of analytic functions.

³Recalling that $|\psi(t, \mathbf{x})|^2$ is the probability density for a particle to be in a volume $d^3\mathbf{x}$ around $\mathbf{x}$.

⁴A representation T of a group G is a homomorphism from G to a group $\mathcal{L}(V)$ composed of linear operators, which exert their action over a vector space V . T being a homomorphism implies that the image Γ of G through T , $\Gamma = \{T(G)\}$, is itself a group, specifically a subgroup of $\mathcal{L}(V)$. Strictly speaking, a representation T shows how the elements of the group G would act on a vector space V if they were linear operators. Without rigour, Γ is often referred to as the representation of G , instead of the homomorphism T .

The *fundamental* representation of the Lorentz group is made of four-dimensional matrices acting on the Minkowski space-time. By definition, when a Lorentz transformation Λ is performed on x , a quadri-vector is such that $A^{\mu'}(x') = \Lambda^\mu{}_{\nu'} A^\nu(x)$. One can therefore find that the operator that gives $A^{\mu'}(x)$ from $A^\mu(x)$, when a Lorentz transformation on x is performed, can be actually interpreted as the angular momentum operator with $s = 1$.

Taking 1.11 into account, one can easily verify that the locally invariant Lagrangian $\mathcal{L}$ can be written as

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\mathcal{D}_\mu\psi - m\bar{\psi}\psi \quad (1.13)$$

where

$$\mathcal{D}_\mu = \partial_\mu + iqA_\mu. \quad (1.14)$$

Equation 1.14 is the so-called minimal coupling which is introduced in classical electrodynamics to describe charged particles in an electromagnetic field. We can therefore conclude that the gauge field A_μ is the quantum field associated with the quantum force-carrier of the electromagnetic force, the mass-less photon.

What has just been displayed is the principle of local gauge invariance, which can be summarized as the following: the request that $\mathcal{L}$ be locally invariant under a symmetry group makes it mandatory to introduce additional **mass-less boson fields**, the gauge fields, obeying appropriate transformation laws under the symmetry group. Furthermore, the same request determines the *form* of the interaction between the initial free fields and the gauge fields, this coupling always being of current-field type.

As we have shown, the symmetry group for Quantum Electrodynamics (QED) is U(1), with A_μ being the quantum field describing the photon.

To ensure that A_μ is part of the dynamical system, gauge-invariant terms built upon A_μ itself and its derivatives must be included in the total Lagrangian $\mathcal{L}$. Adding the kinetic term $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$, with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ being the electromagnetic tensor, meets this request. It is noteworthy that $F_{\mu\nu}$ satisfies the relation (which shall be later generalized):

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = iqF_{\mu\nu}. \quad (1.15)$$

In conclusion, the locally invariant Lagrangian is

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\mathcal{D}_\mu\psi - m\bar{\psi}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (1.16)$$

1.2.3 Local invariance of the Lagrangian under SU(N)

Since the local gauge invariance principle turned out to be extremely successful in describing the QED, a similar approach involving non-abelian Lie groups was pursued by Chen Ning Yang and Robert Mills in the early '50s in order to describe the remaining forces. In particular, the Yang-Mills theory focuses on SU(N), a class of non-abelian Lie groups of dimension $n = N^2 - 1$ consisting of unitary transformations such that their determinant equals +1 [76]. In 1972 Gerard 't Hooft proved Yang-Mills theory to be renormalizable⁵ [2].

As a non-abelian group, SU(N) is such that its generators⁶ $\mathbf{T} = \{T_i\}_{i=1}^n$ (and, therefore, all its elements) do not commute with each other; instead, they satisfy a commutation relation:

$$[T_i, T_j] = if_{ijk}T_k \quad (1.17)$$

⁵A renormalizable theory is such that, when equations are written in terms of the *physical* parameters instead of the *bare* parameters (i.e. those included so far in the Lagrangian, such as m and q) the predicted quantities are finite.

⁶Recalling that each element of a continuous Lie group can be seen as an analytic function of n parameters, the generators $\{T_i\}_{i=1}^n$ of the group are elements such that:

$$g(\epsilon) \approx \mathbb{1} + \epsilon_i \frac{\partial g}{\partial \epsilon_i} \Big|_{\epsilon=0} \equiv \mathbb{1} + \epsilon_i T_i$$

where $\epsilon = \boldsymbol{\theta}/m$ is an infinitesimal parameter vector ($m \rightarrow \infty$). The number of generators always equals the dimension of the group. By means of the exponential map, a finite transformation (i.e. a succession of an infinite number of infinitesimal transformations) can be written in terms of the generators:

$$g(\boldsymbol{\theta}) = \lim_{m \rightarrow \infty} \left(\mathbb{1} + \frac{\boldsymbol{\theta} \cdot \mathbf{T}}{m} \right)^m = e^{\boldsymbol{\theta} \cdot \mathbf{T}}$$

The generator of U(1) is the quantity which defines the transformations alongside with the parameter $i\theta$. Referring to equation 1.6, the generator of U(1) is therefore $q\mathbb{1}$.

where f_{ijk} are the so-called structure constants of the group. For the elements of the group to be unitary, the generators need to be hermitian: $T_i^\dagger = T_i$. An infinitesimal local transformation U of $SU(N)$ is conventionally written as

$$U(\epsilon(x)) = \mathbb{1} - ig\epsilon_i(x)T_i = \mathbb{1} - ig\epsilon(x) \cdot \mathbf{T} \quad (1.18)$$

where g is called a *coupling constant*.

Suppose we have a number D of Dirac spinor fields with the same mass m : $\psi = (\psi_1, \dots, \psi_D)^\top$. The corresponding free Lagrangian is:

$$\mathcal{L}_0 = i\bar{\psi}_a \gamma^\mu \partial_\mu \psi_a - m\bar{\psi}_a \psi_a = i\bar{\psi} \gamma^\mu \partial_\mu \psi - m\bar{\psi} \psi. \quad (1.19)$$

In order to evaluate the action of $SU(N)$ on the *multiplet* ψ , we choose a D -dimensional representation of $SU(N)$ ⁷, which means that U is a $D \times D$ matrix depending on n parameters. The action of U on $\mathcal{L}_0$ leads to

$$\mathcal{L}'_0 = \mathcal{L}_0 + \bar{\psi} i \gamma^\mu (U^\dagger \partial_\mu U) \psi. \quad (1.20)$$

One can therefore define a new Lagrangian $\mathcal{L}$

$$\mathcal{L} = \mathcal{L}_0 - g\bar{\psi} \gamma^\mu T_i \psi W_\mu^{(i)} \equiv \mathcal{L}_0 - j^{(i)\mu} W_\mu^{(i)} = \mathcal{L}_0 - \mathbf{j}^\mu \cdot \mathbf{W}_\mu \quad (1.21)$$

where $\mathbf{W}_\mu = \{W_\mu^{(i)}\}_{i=1}^n$ are n gauge fields which are introduced to ensure the invariance of $\mathcal{L}$ under infinitesimal local transformations of $SU(N)$ like 1.18:

$$\mathcal{L}' = \mathcal{L}_0 + \bar{\psi} i \gamma^\mu (U^\dagger \partial_\mu U) \psi - g\bar{\psi}' \gamma^\mu T_i \psi' W_\mu^{(i)} = \mathcal{L} = \mathcal{L}_0 - g\bar{\psi} \gamma^\mu T_i \psi W_\mu^{(i)}. \quad (1.22)$$

It can be shown that the request 1.22 is fulfilled when:

$$W_\mu'^{(k)} = W_\mu^{(k)} + \partial_\mu \epsilon_k - g f_{ijk} \epsilon_i W_\mu^{(j)} \quad (1.23)$$

which is the non-abelian counterpart of condition 1.12. The existence of the last term of equation 1.23 is a consequence of the non-abelian structure of the group.

After introducing the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu + ig\mathbf{T} \cdot \mathbf{W}_\mu \quad (1.24)$$

the locally invariant Lagrangian can be again written as in equation 1.13. Generalizing equation 1.15, one can define n tensors $\mathbf{G}_{\mu\nu} = \{G_{\mu\nu}^{(i)}\}_{i=1}^n$ such that

$$[\mathcal{D}_\mu, \mathcal{D}_\nu] = igG_{\mu\nu}^{(i)} T_i = ig\mathbf{G}_{\mu\nu} \cdot \mathbf{T}. \quad (1.25)$$

It can be shown that such tensors can be written as

$$G_{\mu\nu}^{(i)} = \partial_\mu W_\nu^{(i)} - \partial_\nu W_\mu^{(i)} - g f_{lki} W_\mu^{(l)} W_\nu^{(k)} \quad (1.26)$$

so that, finally, the locally invariant Lagrangian turns out to be

$$\mathcal{L} = i\bar{\psi} \gamma^\mu \mathcal{D}_\mu \psi - m\bar{\psi} \psi - \frac{1}{4} \mathbf{G}_{\mu\nu} \cdot \mathbf{G}^{\mu\nu}. \quad (1.27)$$

The last term of 1.26 allows for interactions between the gauge bosons associated with the gauge fields $W_\mu^{(i)}$. The latter are referred to as Yang-Mills fields and are always associated with mass-less gauge bosons with $s = 1$; moreover, the term which describes the gauge boson-particle coupling is always of a current-field type, as one can infer by looking at equation 1.21.

As in the abelian case, the existence of $W_\mu^{(i)}$ ensures the local invariance of $\mathcal{L}$, whereas the currents $j_\mu^{(i)}$ arise as conserved quantities from the global gauge invariance.

⁷If the representation has dimension N , it is called the fundamental representation.

1.3 The Standard Model

Bearing in mind the local gauge invariance principle as a starting point, the Standard Model can be built on the following steps:

- a symmetry group G is chosen;
- the fields describing the particles are picked, alongside with their representations under G (i.e. the way the fields transform under G);
- the free Lagrangian of the fields is written;
- the request for the free Lagrangian being locally invariant under G gives rise to as many mass-less gauge fields as the number of generators of G ;
- the Brout–Englert–Higgs mechanism (see section 1.3.2) is applied in order to take the masses of some gauge bosons into account in a gauge-invariant way.

The symmetry group of the Standard Model is $G = \text{SU}(2)_L \otimes \text{U}(1)_Y \otimes \text{SU}(3)_c$. In particular:

- $\text{SU}(2)_L \otimes \text{U}(1)_Y$ allows for the weak and the electromagnetic interactions;
- $\text{SU}(3)_c$ is responsible for the strong interaction.

In the following section we will get acquainted with the fundamentals of the electroweak and the strong interactions, alongside with the Brout–Englert–Higgs mechanism and the model behind fermions masses.

1.3.1 The Electroweak theory

The electroweak theory consists in the electromagnetic and the weak forces unified description, which was firstly provided by Sheldon Glashow in 1961 [55]. Later on, Steven Weinberg [75] and Abdus Salam [71] integrated the Brout–Englert–Higgs mechanism in Glashow’s theory, giving birth to the modern GWS model.

In $\text{SU}(2)_L \otimes \text{U}(1)_Y$, $\text{SU}(2)_L$ is the symmetry group of the weak interaction. The subscript L is related to a symmetry called *chiral*: two chiral projection operators are defined, namely the left-handed and the right-handed chiral operators:

$$P_L = \frac{1}{2}(1 - \gamma^5), \quad P_R = \frac{1}{2}(1 + \gamma^5). \quad (1.28)$$

Since $P_L + P_R = \mathbb{1}$, any fermion field can be decomposed into the sum of its left-handed and right-handed components: $f = P_L f + P_R f = f_L + f_R$. The fermions fields actually interacting via the *charged* weak force carriers are f_L , whereas both f_L and f_R are involved with the neutral current weak processes (i.e. those mediated by Z^0).

In the GWS model, the chosen representation of $\text{SU}(2)$ is the fundamental one, so that the generators are represented by the 2×2 trace-less and hermitian Pauli matrices $\{\sigma_1, \sigma_2, \sigma_3\}$:

$$T_i = \frac{\sigma_i}{2}, \quad [\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \quad (1.29)$$

where the Levi-Civita symbol ε_{ijk} provides $\text{SU}(2)$ with its structure constants. In the GWS model, the generators are referred to as weak *isospin* operators I_i and the eigenvalues of I_3 give the possible values of the quantum number called weak isospin. Given the bi-dimensional representation of $\text{SU}(2)_L$, the left-handed components of the fermion fields are organized into doublets which undergo $\text{SU}(2)_L$ transformations, while the right-handed components remain unchanged and are therefore arranged into weak isospin singlets:

$$\text{SU}(2)_L : \quad \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}'_L = e^{-ig\frac{\sigma}{2} \cdot \theta(x)} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}_L, \quad f'_R = f_R. \quad (1.30)$$

The opposite categorization applies to the antifermions fields: the right-handed components are doublets of $\text{SU}(2)_L$, whereas the left-handed components are singlets. A summary of $\text{SU}(2)_L$ doublets and singlets is

given in Table 1.1. The down-like quark fields actually involved in weak processes are those obtained from the quark mass eigenstates by means of the Cabbibo-Kobayashi-Maskawa matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.31)$$

Moreover, as far as the Standard Model is concerned, neutrinos are described by the left-handed components only (antineutrinos are described by right-handed fields only).

	Generation			Quantum numbers			
	I	II	III	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	0
	e^-_R	μ^-_R	τ^-_R	1/2	-1/2	-1	-1
	$\begin{pmatrix} e^+ \\ \bar{\nu}_e \end{pmatrix}_R$	$\begin{pmatrix} \mu^+ \\ \bar{\nu}_\mu \end{pmatrix}_R$	$\begin{pmatrix} \tau^+ \\ \bar{\nu}_\tau \end{pmatrix}_R$	0	0	-2	-1
Antileptons	$\begin{pmatrix} e^+ \\ \bar{\nu}_e \end{pmatrix}_R$	$\begin{pmatrix} \mu^+ \\ \bar{\nu}_\mu \end{pmatrix}_R$	$\begin{pmatrix} \tau^+ \\ \bar{\nu}_\tau \end{pmatrix}_R$	1/2	+1/2	+1	+1
	e^+_L	μ^+_L	τ^+_L	1/2	-1/2	+1	0
	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	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	+1/3	+2/3
	u_R	c_R	t_R	1/2	-1/2	+1/3	-1/3
	d'_R	s'_R	b'_R	0	0	+4/3	+2/3
Antiquarks	$\begin{pmatrix} \bar{u} \\ \bar{d}' \end{pmatrix}_R$	$\begin{pmatrix} \bar{c} \\ \bar{s}' \end{pmatrix}_R$	$\begin{pmatrix} \bar{t} \\ \bar{b}' \end{pmatrix}_R$	0	0	-2/3	-1/3
	$\bar{u}_L$	$\bar{c}_L$	$\bar{t}_L$	1/2	+1/2	-1/3	+1/3
	$\bar{d}'_L$	$\bar{s}'_L$	$\bar{b}'_L$	1/2	-1/2	-1/3	-2/3
	$\bar{u}_L$	$\bar{c}_L$	$\bar{t}_L$	0	0	-4/3	-2/3
	$\bar{d}'_L$	$\bar{s}'_L$	$\bar{b}'_L$	0	0	+2/3	+1/3

Table 1.1: Overview of the quantum numbers of the Standard Model fermions in the GWS model. The right-handed neutrinos do not participate in the SM interactions and they are therefore not considered here.

As regards $U(1)_Y$, the subscript Y stands for the weak *ipercharge* operator, which is the generator of the group and whose combination with I_3 by means of the Gell-Mann-Nishijima formula

$$I_3 + \frac{Y}{2} = Q \quad (1.32)$$

gives rise to the generator Q of the electromagnetic interaction symmetry group $U(1)_{\text{em}}$. By way of explanation, in order to consider the weak and the electromagnetic interaction as one, a non-trivial combination of $SU(2)_L$ and $U(1)_{\text{em}}$ generators must be taken into account.

Both $SU(2)_L$ doublets and singlets undergo $U(1)_Y$ transformations:

$$U(1)_Y : \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}'_L = e^{-ig'Y\delta(x)} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}_L, \quad f'_R = e^{-ig'Y\delta(x)} f_R. \quad (1.33)$$

It is noteworthy that $SU(2)_L$ and $U(1)_Y$ have different coupling constants ($g \neq g'$).

Temporarily neglecting the mass terms of the fermions (see section 1.3.4), the free Lagrangian is

$$\mathcal{L} = i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + i\bar{\psi}_R \gamma^\mu \partial_\mu \psi_R + \text{h.c.} \quad (1.34)$$

where:

- ψ_L and ψ_R are a left-handed doublet and a right-handed singlet of $SU(2)_L$, respectively;
- the sum over all the fermions is implied;

- h.c. stands for the hermitian conjugate and takes account of the antifermions.

Since $SU(2)_L \otimes U(1)_Y$ is the direct product of two different groups, we can define two separate covariant derivatives to make $\mathcal{L}$ locally invariant:

$$\mathcal{D}_\mu^L = \partial_\mu + ig \frac{\boldsymbol{\sigma}}{2} \cdot \mathbf{W}_\mu + ig' Y B_\mu, \quad \mathcal{D}_\mu^R = \partial_\mu + ig' Y B_\mu \quad (1.35)$$

where $\mathcal{D}_\mu^L$ and $\mathcal{D}_\mu^R$ act on doublets and singlets, respectively. While the gauge fields $\mathbf{W}_\mu$ satisfy 1.23 condition with respect to $SU(2)_L$, the gauge field B_μ changes like 1.12 under $U(1)_Y$.

The locally invariant to $SU(2)_L \otimes U(1)_Y$ Lagrangian is then

$$\mathcal{L} = i\bar{\psi}_L \gamma^\mu \mathcal{D}_\mu^L \psi_L + i\bar{\psi}_R \gamma^\mu \mathcal{D}_\mu^R \psi_R - \frac{1}{4} C^{\mu\nu} C_{\mu\nu} - \frac{1}{4} \mathbf{G}^{\mu\nu} \cdot \mathbf{G}_{\mu\nu} \quad (1.36)$$

where $C_{\mu\nu}$ and $\mathbf{G}_{\mu\nu}$ are tensors which satisfy equation 1.26, i.e.:

$$C_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.37)$$

$$G_{\mu\nu}^{(i)} = \partial_\mu W_\nu^{(i)} - \partial_\nu W_\mu^{(i)} - g(\mathbf{W}_\mu \times \mathbf{W}_\nu)^{(i)}. \quad (1.38)$$

Any massive term for the gauge fields would break the local gauge invariance. Nevertheless, the phenomenology of the weak interactions suggests that the weak force carriers have non vanishing masses: here is where the Brout–Englert–Higgs mechanism comes into play.

1.3.2 The Brout–Englert–Higgs mechanism

In 1962, Jeffrey Goldstone, Salam and Weinberg showed [57] that if there is continuous symmetry transformation under which the Lagrangian is invariant, then either the vacuum state is also invariant under the transformation, or there must exist as many spinless particles of zero mass (Goldstone bosons) as the generators of the continuous group. This concept is at the heart of the *spontaneous symmetry breaking*: the fact that the system falls into a ground state that does not exhibit the same symmetry of the Lagrangian results in that very symmetry being hidden or *broken*. Two years later, Peter Higgs [61] and, independently, Robert Brout and Francois Englert [52] applied the spontaneous symmetry breaking concept to a locally invariant to $U(1)$ Lagrangian, showing that the massless gauge boson arising from the gauge principle could acquire mass in a gauge invariant way at the expense of the Goldstone scalar boson.

The asymmetric ground state, on whose existence the spontaneous symmetry breaking lies, can be realised by requiring that the expectation value of some field be non-zero. Let us therefore consider a complex scalar field

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2). \quad (1.39)$$

The Lagrangian of this field is made invariant under $U(1)$ by means of the covariant derivative 1.14:

$$\mathcal{L} = (\mathcal{D}_\mu \phi)^* (\mathcal{D}^\mu \phi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - V(\phi) \quad (1.40)$$

where $V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2$ is the potential term. For the vacuum state of the system to be stable (i.e. for V to exhibit a minimum value), the condition $\lambda > 0$ is essential. If $\mu^2 < 0$, the values of ϕ for which V is minimum define a circle (in the complex plane) of radius v such that

$$v^2 = 2\phi_{\text{vac}}^* \phi_{\text{vac}} = -\frac{\mu^2}{\lambda} \quad (1.41)$$

One can therefore write ϕ_{vac} as

$$\phi_{\text{vac}} = \frac{v}{\sqrt{2}} e^{i\alpha} \quad (1.42)$$

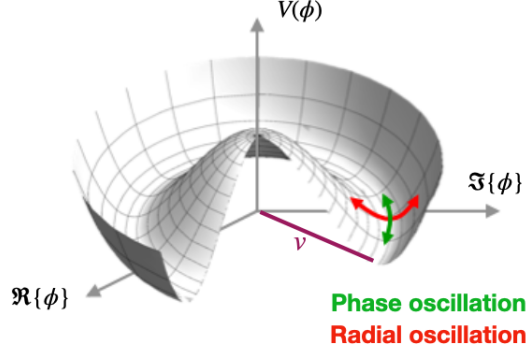


Figure 1.4: Higgs potential for a complex scalar field. The phase (red) and the radial oscillations (green) around the vacuum state are displayed.

with α being an arbitrary phase factor. Whereas $\mathcal{L}$ is invariant to $U(1)$ by definition, ϕ_{vac} is not: the action of a $U(1)$ transformation on 1.42 leads to a different point of the circle. Let us suppose that the vacuum state which the system falls spontaneously into is defined by $\alpha = 0$. One can think of the field ϕ as an expansion around this vacuum state, namely

$$\phi = \frac{v + H(x)}{\sqrt{2}} e^{i\frac{\theta(x)}{v}} \quad (1.43)$$

where $H(x)$ and $\theta(x)$ take account of the radial and phase oscillations around $\phi_{\text{vac}} = v/\sqrt{2}$, respectively (Figure 1.4). Substituting 1.43 in 1.40, one can find that, while the Goldstone boson $\theta(x)$ remain massless, massive terms for both the scalar boson field $H(x)$ and the gauge field A_μ arise, so that the system seems to gain one extra degree of freedom. Actually, the local gauge invariance of $\mathcal{L}$ allows for a $U(1)$ transformation such that

$$\phi' = e^{-iq\beta(x)} \phi = \frac{(v + H(x))}{\sqrt{2}}. \quad (1.44)$$

In doing so, one degree of freedom belonging to the Goldstone boson $\theta(x)$ is *gauged away*, letting the gauge boson A_μ acquire mass while preserving gauge invariance.

After Glashow proposed $SU(2)_L \otimes U(1)_Y$ as the electroweak symmetry group, Weinberg and Salam probed the Brout-Englert-Higgs mechanism in the case of a weak isospin doublet of complex scalar fields with $Y = 1$ (namely, the Higgs field)

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.45)$$

by adding the term

$$\mathcal{L}_H = (\mathcal{D}_\mu^L \phi)^\dagger (\mathcal{D}^{L\mu} \phi) - V(\phi) \quad (1.46)$$

to the Lagrangian 1.36. Here, $\mathcal{D}_\mu^L$ is the covariant derivative 1.35, while the potential $V(\phi)$ is

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{4} (\phi^\dagger \phi)^2. \quad (1.47)$$

Again, picking one quadruplet $\phi_1, \dots, \phi_4$ among the set of all the (physically) equivalent vacuum states, writing the Higgs field as an expansion around the chosen ϕ_{vac} after a gauge transformation such that the massless Goldstone bosons are gauged away, and substituting such a ϕ in the Lagrangian, one finds (as the

gauge fields and the Higgs field are concerned):

$$\begin{aligned}\mathcal{L}_{G\phi} = & \frac{1}{2}\partial_\mu H\partial^\mu H - \lambda v^2 H^2 \\ & - \frac{1}{4}(\partial_\mu W_\nu^{(1)} - \partial_\nu W_\mu^{(1)})(\partial^\mu W^{(1)\nu} - \partial^\nu W^{(1)\mu}) + \frac{1}{8}g^2 v^2 W_\mu^{(1)}W^{(1)\mu} \\ & - \frac{1}{4}(\partial_\mu W_\nu^{(2)} - \partial_\nu W_\mu^{(2)})(\partial^\mu W^{(2)\nu} - \partial^\nu W^{(2)\mu}) + \frac{1}{8}g^2 v^2 W_\mu^{(2)}W^{(2)\mu} \\ & - \frac{1}{4}(\partial_\mu W_\nu^{(3)} - \partial_\nu W_\mu^{(3)})(\partial^\mu W^{(3)\nu} - \partial^\nu W^{(3)\mu}) - \frac{1}{4}C^{\mu\nu}C_{\mu\nu} \\ & + \frac{1}{8}v^2(gW_\mu^{(3)} - g'B_\mu)(gW^{(3)\mu} - g'B^\mu).\end{aligned}$$

The first line refers to the massive Higgs field ($m_H = \sqrt{2\lambda}v$); the second and third lines contain the kinetic and the massive terms of $W_\mu^{(1)}$ and $W_\mu^{(2)}$ ($M_W = gv/2$); the fourth line takes account of $W_\mu^{(3)}$ and B_μ kinetic terms, while the last line is a massive term of a combination of $W_\mu^{(3)}$ and B_μ . Given the Weinberg angle θ_W such that

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (1.48)$$

one can define the gauge fields Z_μ and A_μ

$$Z_\mu = W_\mu^{(3)} \cos\theta_W - B_\mu \sin\theta_W \quad (1.49)$$

$$A_\mu = W_\mu^{(3)} \sin\theta_W + B_\mu \cos\theta_W \quad (1.50)$$

Rearranging the Lagrangian in terms of these fields, A_μ is found to be massless, while Z_μ has mass $m_Z = m_W / \cos\theta_W$; in addition, $W_\mu^{(1)}$ and $W_\mu^{(2)}$ turn out to be mixed into two fields $W_\mu^{(+)}$ and $W_\mu^{(-)}$ (with the same mass as before) such that

$$W_\mu^{(\pm)} = \frac{W_\mu^{(1)} \mp iW_\mu^{(2)}}{\sqrt{2}}. \quad (1.51)$$

Last but not least, in order to obtain the proper electromagnetic covariant derivative, the following condition must be true:

$$q = g \sin\theta_W. \quad (1.52)$$

Equation 1.52 encodes the essence of the electroweak unification.

In summation, the gauge fields responsible for the electroweak interaction are A_μ , $W_\mu^{(+)}$, $W_\mu^{(-)}$ and Z_μ , whose quanta are the massless photon γ and the massive bosons W^+ , W^- and Z^0 , respectively. The existence of the weak neutral current (mediated by Z^0) was proven in 1973 by the Gargamelle collaboration [60]; ten years later, the UA1 [9, 10] and the UA2 [45, 46] collaborations provided the observation of the W and Z bosons at the Super Proton Synchrotron.

In 2012, the ATLAS and CMS Collaboration announced the observation of a particle with a 125 GeV mass compatible with the Standard Model Higgs boson [28, 40]. This discovery earned Higgs and Englert the Nobel Prize in Physics in 2013.

1.3.3 The Quantum Chromodynamics

The QCD is the theory of the strong interaction between quarks and gluons inside the hadrons. It is a non-abelian local gauge theory based on $SU(3)_c$, where the subscript c stands for *color*. The strong charge, also called the color charge, is the QCD counterpart of the electromagnetic charge and was proposed in 1964

by Greenberg [58] and Han-Nambu [59], independently. They theorized three different "values" for the color charge, so that the field representing a quark of flavour f would be

$$\psi^f(x) = \begin{pmatrix} \psi_R^f(x) \\ \psi_G^f(x) \\ \psi_B^f(x) \end{pmatrix} \quad (1.53)$$

with $\psi_i^f(x)$ being Dirac spinors: each quark flavour was theorized to exist into three different color states and SU(3), with its fundamental representation consisting of 3×3 matrices acting on color triplets, was designated as the symmetry group of the strong interaction.

SU(3)_c transformations are generated by the eight Gell-Mann matrices λ_i :

$$\psi_i'^f(x) = e^{ig_s \frac{\lambda}{2} \cdot \theta} \psi_i^f(x) \quad (1.54)$$

where $g_s = \sqrt{4\pi\alpha_s}$ is the strong coupling constant. The local gauge invariance principle gives rise to eight massless gauge fields $A_\mu^{(i)}(x)$, which are directly associated with the strong interaction massless carriers, the gluons. The QCD Lagrangian turns out to be

$$\mathcal{L}_{\text{QCD}} = i\bar{\psi}_i^f(x)\gamma^\mu \mathcal{D}_\mu \psi_i^f(x) - \frac{1}{4} \mathbf{G}^{\mu\nu} \cdot \mathbf{G}_{\mu\nu} \quad (1.55)$$

where

$$\mathcal{D}_\mu = \partial_\mu + ig_s \frac{\lambda}{2} \cdot \mathbf{A}_\mu \quad (1.56)$$

$$G_{\mu\nu}^{(i)} = \partial_\mu A_\nu^{(i)} - \partial_\nu A_\mu^{(i)} - g_s f_{ijk} A_\mu^{(j)} A_\nu^{(k)} \quad (1.57)$$

with f_{ijk} being the SU(3) structure constants.

Since an octet arises from the $3 \otimes \bar{3}$ direct product, each gluon is associated with a color-anticolor combination, so that gluons are said to be bi-coloured: differently from the photon, gluons carry the strong charge. In consequence, they are capable of self-interactions, a feature that is taken account of by the non-abelian nature of the symmetry group, due to which auto-interactions terms arising from 1.57 appear in the QCD Lagrangian.

Studying the quark-antiquark interaction and the three quarks interaction from a perturbation point of view, one can find that the short-distance potential is in both cases attractive, provided that the system is in a singlet state of SU(3)_c. Mesons (quark-antiquark) and barions (three quarks) being SU(3)_c singlets is consistent with the experimental fact that no color charge is observed for composite hadrons. Indeed, QCD exhibits a peculiar property, the so-called *color confinement*, which is itself a consequence of another distinctive feature of the quantum chromo-dynamics: the *asymptotic freedom*.

The *asymptotic freedom* arises from the non-abelian structure of SU(3)_c. Both QED and QCD predictions can be obtained by means of a perturbation approach, so that a given quantity is actually the sum of all the terms of a perturbation series. High order terms of the series are involved with Feynman diagrams containing loops like those in Figure 1.5. The existence of such loops results in the theory providing infinite values of the predictions; fortunately, Yang-Mills theories are proved to be renormalizable, meaning that divergences are taken care of by substituting the bare parameters with physical parameters, including the renormalized coupling constant. In the QCD case, due to the auto-interaction terms of the Lagrangian coming from the non-abelian structure of SU(3)_c (terms that are therefore absent from the QED picture), higher order Feynman diagrams containing quark-antiquark and gluon loops result in the strong coupling constant g_s depending on the distance in such a way that the closer the quarks are, the weaker the intensity of the interaction is, to the point that they can actually be considered *free*. In contrast, distant quarks undergo such an intense attraction that separating two quarks would require an infinite amount of energy. Therefore, as the quark-quark distance increases, the color field creates additional quarks and anti-quarks, which bind to the original quarks in such a way that no free quark can be observed in the final state.

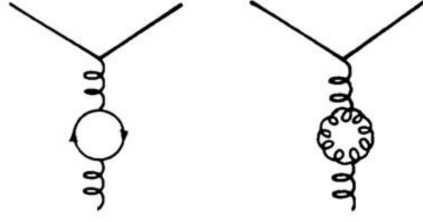


Figure 1.5: QCD Feynman diagrams containing a fermion (left) and a gluon (right) loop. The combination of both amplitudes results in the asymptotic freedom of the QCD.

1.3.4 Fermions masses

Up until now, we have taken fermion masses into account by adding a Dirac mass term $m\psi\bar{\psi}$ to the Lagrangian, like in equation 1.13. This term can actually be written as

$$m\psi\bar{\psi} = m(\psi_L\bar{\psi}_R + \psi_R\bar{\psi}_L). \quad (1.58)$$

Even though equation 1.58 is invariant under $U(1)$, this is not the case for the Standard Model symmetry group $SU(2)_L \otimes U(1)_Y \otimes SU(3)_c$: more precisely, ψ_L and ψ_R undergo different transformations under $SU(2)_L$, so that the Dirac term 1.58 results in not being invariant. In order to take account of fermion masses in a gauge invariant way, an interaction term between the fermion field and the Higgs field can be added to the Lagrangian, so that mass terms arise by means of the spontaneous symmetry breaking mechanism. This additional term is referred to as Yukawa interaction term and has the form

$$\mathcal{L}_{\text{Yukawa}} = -G_L^{ij} \bar{L}_L^i \phi l_R^j - G_d^{ij} \bar{Q}_L^i \phi d_R^j - G_u^{ij} \bar{Q}_L^i \phi_C u_R^j + \text{h.c.} \quad (1.59)$$

where L_L^i and Q_L^i ($i = 1, 2, 3$) are leptons and quarks $SU(2)_L$ doublets (first and fifth row of Table 1.1, respectively), l_R^j , u_R^j , d_R^j ($j = 1, 2, 3$) are leptons and quarks $SU(2)_L$ singlets (second, sixth and seventh row of Table 1.1, respectively), ϕ is the Higgs field 1.45 and ϕ_C is the conjugate $SU(2)_L$ doublet of ϕ , defined as

$$\phi_C = i\sigma_2 \phi^* = \begin{pmatrix} \phi^{0*} \\ -\phi^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_3 - i\phi_4 \\ -(\phi_1 - i\phi_2) \end{pmatrix}. \quad (1.60)$$

As usual, the hermitian conjugate takes account of the anti-fermion terms of the Lagrangian. The 3×3 matrices G_L^{ij} , G_d^{ij} , G_u^{ij} are known as Yukawa matrices: while G_L^{ij} has a diagonal form, G_d^{ij} and G_u^{ij} are non-diagonal in order to take account of the quark mixing. Operating the symmetry breaking mechanism, i.e. choosing a specific vacuum state, writing the Higgs field as expansion terms around ϕ_{vac} and applying a $SU(2)_L$ transformation to ϕ such that the Goldstone bosons are gauged away, the lepton Yukawa term of the Lagrangian turns out to be

$$\mathcal{L}_{\text{Yukawa}}^{\text{lepton}} = - \left(1 + \frac{H}{v} \right) [\bar{l}_L^i m_l^i l_R^i + \bar{l}_R^i m_l^i l_L^i + \text{h.c.}] = -m_l^i \bar{l}^i l^i - \frac{gm_l^i}{2m_W} \bar{l}^i l^i H + \text{h.c.} \quad (1.61)$$

where $m_l^i = G_L^{ii} v / \sqrt{2}$ are the lepton masses, while the quark term is

$$\mathcal{L}_{\text{Yukawa}}^{\text{quark}} = - \left(1 + \frac{H}{v} \right) [\bar{u}_L^i M_u^{ij} u_R^j + \bar{d}_L^i M_d^{ij} d_R^j + \text{h.c.}] \quad (1.62)$$

where

$$M_d^{ij} = G_d^{ij} \frac{v}{\sqrt{2}}, \quad M_u^{ij} = G_u^{ij} \frac{v}{\sqrt{2}} \quad (1.63)$$

are non-diagonal matrices that need to be diagonalized in order to write the Lagrangian in terms of the quark mass fields. The diagonalization can be carried out by means of unitary matrices U_u^L , U_u^R , U_d^L , U_d^R . In summation, the Yukawa term for all the fermions is

$$\mathcal{L}_{\text{Yukawa}} = -m_l^i \bar{l}^i l^i - \frac{gm_l^i}{2m_W} \bar{l}^i l^i H - m_d^\alpha \bar{d}^\alpha d^\alpha - \frac{gm_d^\alpha}{2m_W} \bar{d}^\alpha d^\alpha H - m_u^\alpha \bar{u}^\alpha u^\alpha - \frac{gm_u^\alpha}{2m_W} \bar{u}^\alpha u^\alpha H + \text{h.c.} \quad (1.64)$$

where d^α and u^α ($\alpha = 1, 2, 3$) are mass eigenstates that are connected to the weak interaction eigenstates u^i , d^i :

$$u_L^\alpha = (U_u^L)^{\alpha i} u_L^i \quad u_R^\alpha = (U_u^R)^{\alpha i} u_R^i \quad (1.65)$$

$$d_L^\alpha = (U_d^L)^{\alpha i} d_L^i \quad d_R^\alpha = (U_d^R)^{\alpha i} d_R^i \quad (1.66)$$

and m_d^α , m_u^α are the eigenvalues of M_d^{ij} and M_u^{ij} , respectively. The U matrices are involved with the CKM matrix, which appears in the charged weak currents for the quarks:

$$V^{\alpha\beta} = \left[U_u^L U_d^{L\dagger} \right]^{\alpha\beta}. \quad (1.67)$$

As it is made clear from equation 1.64, Dirac mass terms for both leptons and quarks are obtained in a gauge invariant way, as it should be. In addition, the spontaneous symmetry breaking gives rise to interaction terms that couple the Higgs field to the fermion fields, these terms being proportional to the fermion masses. Since the G matrices eigenvalues are not fixed by the theory, the Standard Model does not predict the values of the fermion masses.

As the right-handed neutrino fields do not participate in the Standard Model interactions, the corresponding Yukawa term is vanishing, so that the SM neutrinos are rigorously massless.

1.3.5 Vector Boson Scattering

As already pointed out, auto-interaction terms between the weak gauge bosons are accounted for by terms of the Standard Model Lagrangian arising from Equation 1.26. Consequently, the scattering of vector bosons (VBS, Figure 1.6) is a process predicted by the Standard Model and represents a key tool for probing the non-Abelian gauge structure of the electroweak interaction.

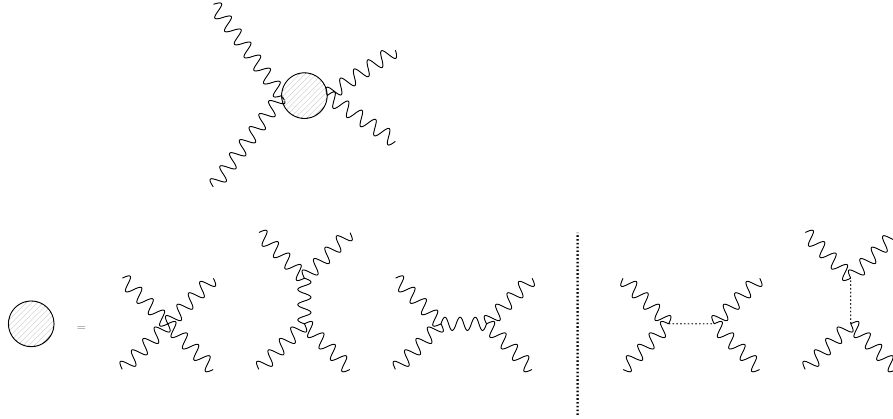


Figure 1.6: Feynman diagrams of the vector boson scattering. Both vector bosons auto-interactions diagrams and diagrams involving vector bosons couplings with the Higgs boson (dashed line) contribute to the SM VBS amplitude, at tree-level.

Moreover, VBS is intimately connected to the Higgs boson: the scattering amplitude of longitudinal polarized vector bosons needs a contribution from the exchange of the SM Higgs boson in order to preserve unitarity at a centre-of-mass energy of order 1 TeV [68, 69]. The discovery of a scalar boson in 2012 at the LHC, with measured properties [26, 34, 41, 47] consistent with those predicted for the SM Higgs boson, represents a major breakthrough in the comprehension of the electroweak symmetry breaking, and the study of VBS can help further investigate its nature. In fact, any deviation in the Higgs boson coupling with the vector bosons predicted by the Standard Model (*anomalous* couplings) would disrupt the balance between the amplitudes contributing to the VBS cross section. If this is the case, studying VBS processes allows us to investigate whether the SM Higgs mechanism is the sole source of electroweak symmetry breaking.

The above mentioned anomalous couplings (in particular, anomalous *quartic* gauge couplings, aQGCs) can be accounted for by means of an effective field theory approach [50]. When searching for physics

beyond standard model, one can either directly search for unknown resonances or look for new interactions between known particles of the standard model. An effective field theory adopts the latter approach in a model-independent way. In addition, the SM gauge symmetries are respected and it is possible to recover the Standard Model in an appropriate limit.

More specifically, in the SM Lagrangian all operators (i.e. the product of fields) are constrained to be of mass dimension four or less⁸. The theory extension is obtained by adding operators of higher dimensions, so that the effective Lagrangian can be written as [51]:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_{n=5}^{\infty} \frac{f_n}{\Lambda^{n-4}} O_n \quad (1.68)$$

where $\mathcal{L}_{\text{SM}}$ is the Standard Model Lagrangian and f_n are dimensionless coupling-strength coefficients which are typically of $\mathcal{O} \sim 1$, while Λ is the mass scale of new physics. By definition, the coefficients by which the additional operators are multiplied are negligible if the mass Λ is large with respect to the experimentally accessible energies. In other words, the operators which contribute the most to the effective Lagrangian are those with the lowest dimensionality.

According to [51], there exist three types of dimension-8 operators that modify the quartic couplings of the gauge bosons:

- scalar operators $O_{S,i}$, $i = 0, \dots, 2$;
- tensor operators $O_{T,i}$, $i = 0, \dots, 9$;
- mixed operators $O_{M,i}$, $i = 0, \dots, 7$.

The resulting effective Lagrangian yields to an enhancement of the SM VBS cross-section at high transverse momentum of the vector bosons and high invariant mass of the diboson system.

My Thesis work is part of an electroweak diboson production search currently performed by the ATLAS Collaboration, which will be illustrated in Chapter 4.

⁸In natural units ($\hbar = c = 1$), the Lagrangian L is of mass dimension one. Therefore, the Lagrangian density $\mathcal{L}$ is of mass dimension four, since $[m]^{-1}$ is of length dimension one.

Chapter 2

ATLAS Experiment at the LHC, CERN

The European Organization for Nuclear Research (in french: Conseil Européen pour la Recherche Nucléaire), known as CERN, is a European research organization that runs the world's largest particle physics laboratory, based in a north-west suburb of Geneva on the Franco-Swiss border. Born from a handful of visionary scientists' desire for unity in post-war Europe to be sealed by a world-class physics research facility, CERN was officially established in 1954 by a convention made of 12 countries in Western Europe. When CERN project was conceived, pure physics research was mainly devoted to the understanding of the atom, hence the word "nuclear" in the organization name. At present, CERN is made of person-power from 23 different states and its main area of research is sub-atomic particle physics. Among the most important achievements made by scientists through experiments at CERN we can cite the discovery of the W , Z and Higgs boson [9, 10, 28, 40, 45, 46, 60], the discovery of direct CP violation [44] and the foundation of the World Wide Web.

2.1 The Large Hadron Collider

At CERN, physicists and engineers from all over the world study the structure of the universe by means of the largest and most complex particle accelerator ever built: the Large Hadron Collider (LHC). Two high-energy proton beams are made to collide inside the accelerator, so that scientists can investigate the physics processes involved with the collisions and try to answer the fundamental question of what the universe is made of, pushing the human knowledge frontier.

The acceleration is achieved by means of a proton source followed by a succession of machines that sequentially accelerate particles to increasingly higher energies, the last element of the chain being the LHC. Here, the protons are accelerated up to the record energy of 6.5 TeV per beam.

Figure 2.1 shows a schematic view of the whole accelerating structure. The proton source is a bottle of hydrogen gas. Hydrogen atoms are stripped of their electrons by means of an electric field, leaving hydrogen nucleus, i.e. protons, behind. The first accelerating machine of the chain is **Linac2**, i.e. a linear accelerator that pushes protons to a 50 MeV energy. The **Proton Synchrotron Booster** and sequentially the **Proton Synchrotron** (PS) accelerate the protons to 1.4 GeV and 25 GeV, respectively. The beam is then injected into the **Super Proton Synchrotron** (SPS) to be accelerated to 450 GeV. Most of these accelerators have their own halls where experiments exploit the beams at lower energies with respect to the energy in the LHC, where the protons are finally transferred.

The LHC beam pipes consist of a 27-kilometres ring of superconducting magnets equipped with numerous accelerating structures that boost the particles along their way around the collider. The magnets are needed to produce an intense magnetic field which keeps the particles on a circular trajectory inside the ring. The protons are split into two beams travelling in opposite directions in such a way that several collision points are found along the ring. Four detectors are installed near the latter points: **ALICE**, **ATLAS**, **CMS** and **LHCb**. The center of mass energy at each collision point equals 13 TeV.

Alongside with protons, other particles are accelerated in the LHC. For instance, lead ions enter **Linac3** from a source of vaporised lead and are then collected and accelerated in the **Low Energy Ion Ring**. Successively, they follow the same route to their maximum energy as the protons.

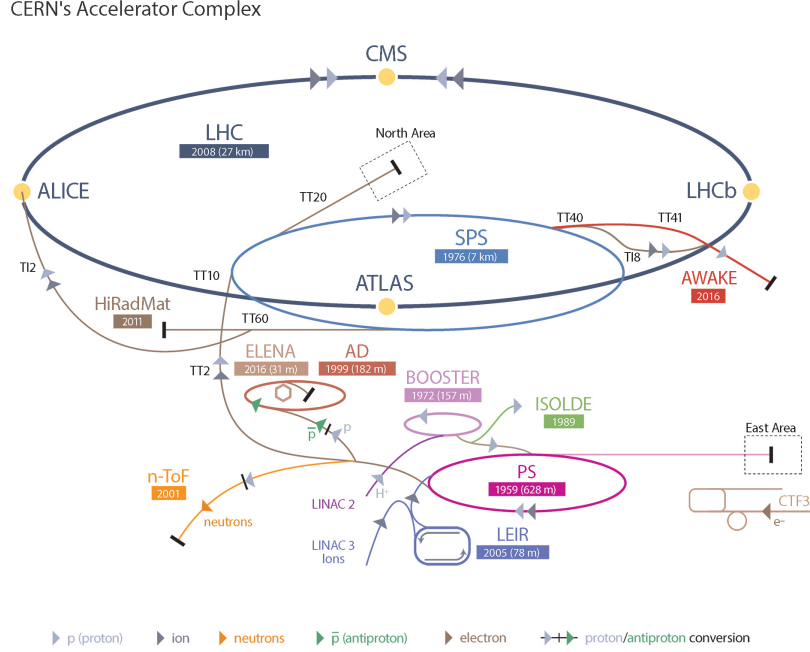


Figure 2.1: Schematic representation of CERN accelerators and experiments installed alongside the accelerators.

2.1.1 Run 1

A construction parameter that is crucial to the discovering power of a collider is its *instantaneous luminosity*, defined as

$$\mathcal{L} = \frac{N_p^2 f k}{4\pi R^2} \quad (2.1)$$

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 in the plane orthogonal to the beam direction. The instantaneous luminosity $\mathcal{L}$ depends only on beam parameters. For a gaussian-profile beam, it is connected to the rate R of events produced by collisions through the cross-section σ of the considered process:

$$R = \mathcal{L} \cdot \sigma. \quad (2.2)$$

Integrating the rate for a given time frame, the proportional factor between the number N of collected events and the cross-section for a given period T is the *integrated luminosity* L :

$$N = L \cdot \sigma, \quad L = \int_T \mathcal{L} dt. \quad (2.3)$$

The first successful circulation of the beams occurred on 10th September 2008. Unfortunately, a serious outage damaged some of the superconducting magnets soon after, causing the LHC to stop its activity until November 2009, after a long technical intervention. First collisions took place on 30th March 2010, with the rest of the year mainly devoted to commissioning.

2011 was the first production year, with an integrated luminosity of 5 fb⁻¹ delivered to both ATLAS and CMS. By the time of the 2012 summer conferences, data corresponding to 6 fb⁻¹ were collected, allowing for the announcement of the discovery of a Higgs-like particle on 4th July 2012.

During the first period (2010-2011), the centre-of-mass energy of the pp system was $\sqrt{s} = 7$ TeV. In 2012, $\sqrt{s}$ was increased to 8 TeV.

In 2011, LHC instantaneous luminosity reached the value of $3.65 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ at its maximum, the design peak luminosity being $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Such a high luminosity was obtained through 1380 (2808 by

design) bunches per beam, each of them containing 10^{11} protons. The bunches have very small transverse spread, about $15\text{ }\mu\text{m}$ in the transverse direction, while the longitudinal length is about 7 cm . The original LHC design planned the bunches to cross every 25 ns , corresponding to a collision rate of 40 MHz . However, the actual bunch time spacing reached in 2011 and 2012 was 50 ns . The resulting 2011-2012 integrated luminosity is shown in Figure 2.2.

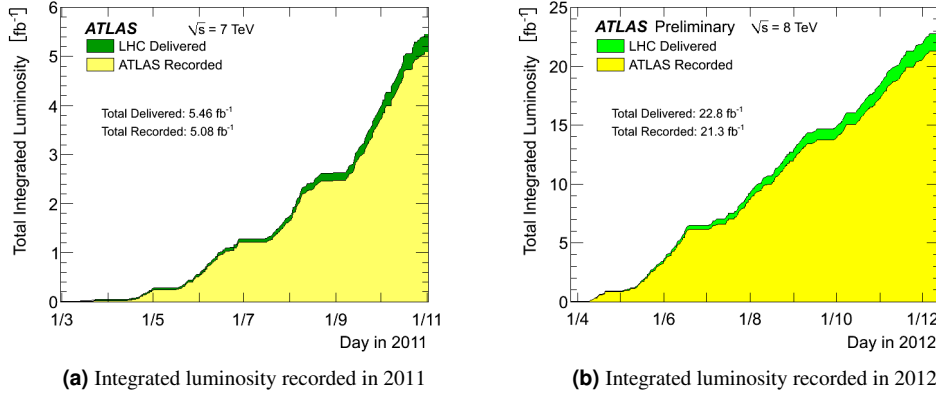


Figure 2.2: The integrated luminosity as a function of time delivered by LHC (green) and recorded by ATLAS (yellow) in 2011 and 2012.

The LHC Run 1 officially ended in 2013.

2.1.2 Run 2

Long Shut-down 1 (LS1) lasted from 2013 to 2014. In May 2015, a second phase of data collection started: the LHC Run 2. The goal of Run 2 consisted in collecting about 100 fb^{-1} of data, at a collision energy of 14 TeV , until 2018.

The initial phase of data collection was dedicated to put the collider into operation, namely by testing the magnets performance and the alignment of the spectrometer, using a beam bunch spacing of 50 ns for 1 fb^{-1} integrated luminosity.

After the commissioning phase, the bunch-spacing was improved to 25 ns and the beams produced collisions at $\sqrt{s} = 13\text{ TeV}$ with a peak luminosity of $5.0 \times 10^{33}\text{ cm}^{-2}\text{ s}^{-1}$.

After the Run 2 first technical stop in 2016, LHC reached a peak luminosity of $12.1 \times 10^{33}\text{ cm}^{-2}\text{ s}^{-1}$. The number of collisions recorded by ATLAS and CMS during the proton run from April to the end of October was 60% higher than anticipated. More precisely, the integrated luminosity received by ATLAS and CMS reached 40 fb^{-1} , compared with the 25 fb^{-1} originally planned.

During the Extended Year End Technical Stop (from the end of 2016 until March 2017), CERN's accelerators took a long break. Thanks to the improvements effected during this period, the instantaneous luminosity doubled its nominal value in 2017, reaching $2.09 \times 10^{34}\text{ cm}^{-2}\text{ s}^{-1}$ (Figure 2.3c).

In the same year, the LHC has provided its two major experiments, ATLAS and CMS, with 50 fb^{-1} of data, i.e. 5 billion million collisions, regardless of a serious setback due to a vacuum problem in the beam pipe of a magnet cell. In fact, even though the number of bunches that could circulate in the machine was restrained, the arrangement of the bunches in the beams was changed in such a way that luminosity started to increase again after a few weeks. The operating parameters were optimised as well, alongside with the size of the beams, which was significantly reduced. As a result, up to 60 collisions were produced at each crossing in 2017.

In 2018 the target integrated luminosity of 60 fb^{-1} for the ATLAS and CMS proton run was exceeded by 10%, resulting in a total integrated luminosity during Run 2 (2015 - 2018) of 140 fb^{-1} , and of 139 fb^{-1} since the start of LHC physics (Figure 2.4). Figure 2.3 shows the peak and the integrated luminosity delivered during each year of Run 2, which officially ended on 3rd December.

LHC Run 2 provided the scientific community with relevant precision measurements, from Higgs boson

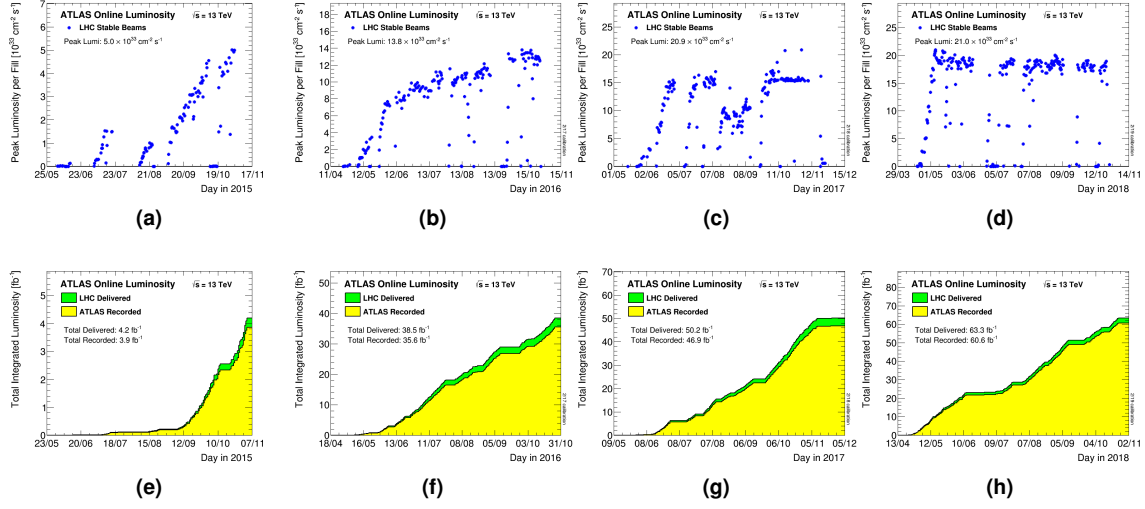


Figure 2.3: The LHC luminosity delivered during Run 2: peak luminosity (first row) and integrated luminosity (second row) from 2015 to 2018 (left to right). In the bottom plot, the total delivered data is shown in green, while the data recorded by ATLAS is shown in yellow.

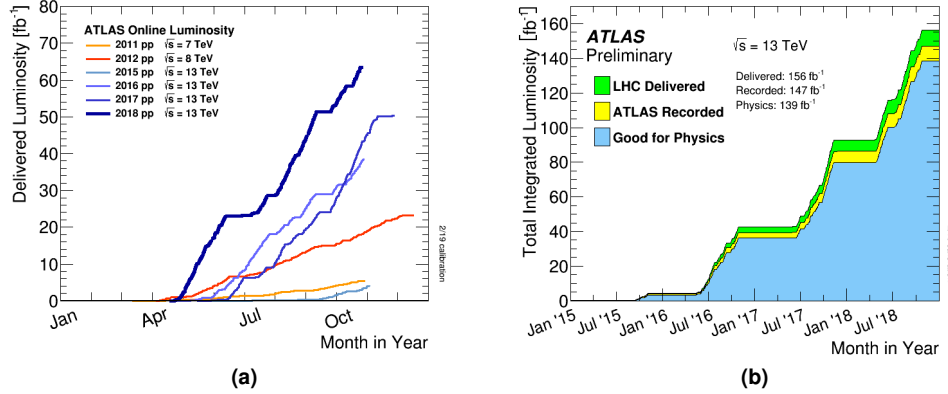


Figure 2.4: (a): LHC Run 2 integrated luminosity for different years of the run. (b): Full Run 2 integrated luminosity delivered by LHC (green), recorder by ATLAS (yellow) and exploited for physics measurements (azure).

mass to top quark and vector boson masses. Precision measurements are crucial to constrain Standard Model parameters that are not predicted by the model.

2.1.3 Run 3 and High Luminosity LHC

The second Long Shut-down (LS2) started on December 2018. It was scheduled to last about two years, until the beginning of 2021. During this time new electronics, trigger and detector components will be installed, with plans to start Run 3 and, successively, guide the LHC towards its High Luminosity phase.

The High Luminosity LHC (HL-LHC) is an upgrade of the LHC scheduled to run from 2027 and designed to achieve instantaneous luminosities five times larger than the LHC nominal value, thereby enabling the experiments to enlarge their data sample by one order of magnitude compared with the LHC baseline programme. As for the ATLAS detector, the HL-LHC will deliver a luminosity seven time larger than the value for which ATLAS was originally designed. In order to accomplish this ambitious task, the development and installation of new detectors with radiation-hard elements, finer granularity and faster readout are needed. While most of these new systems will be installed during the next Long Shutdown (LS3, scheduled for 2025), some will already see service in Run 3, such as the ATLAS 10 m-diameter New Small Wheels (Figure 2.5).

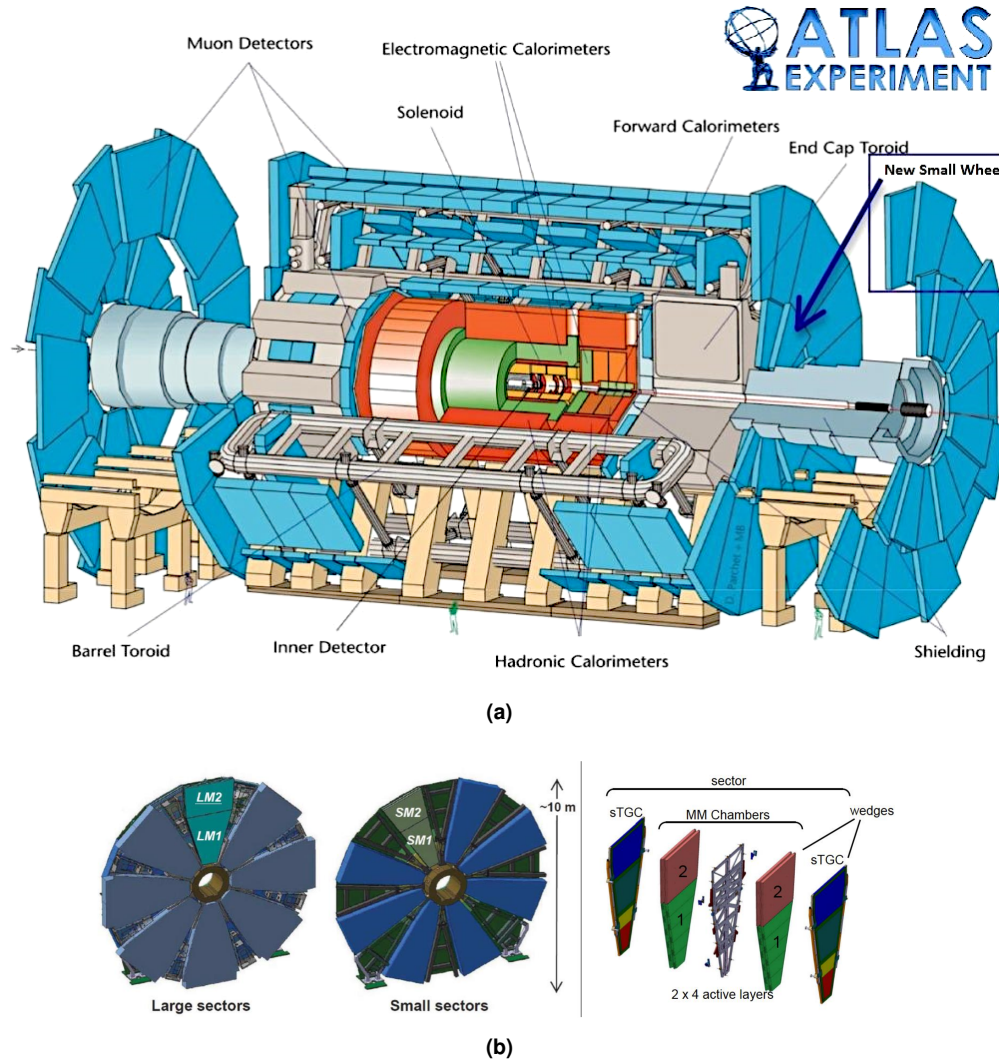


Figure 2.5: (a): Schematic representation of the ATLAS detector. (b): Overall view of the New Small Wheel. Left: Views of the wheel, highlighting the modules large sectors LM1 and LM2 and the modules of the small ones SM1 and SM2. Right: Exploded view of a Large sector, with two sTGC and two MM quadruplets on both sides of the spacer frame.

ATLAS New Small Wheels installation represents the largest and most critical project to be carried out during LS2, and the first major HL-LHC addition to the detector. Teams are currently finalising the construction of the new wheels. The installation should be finalised by the beginning of 2022, but all activities scheduled before 2022 are suffering from delays due to the COVID-19 pandemic.

The New Small Wheels employ two detector technologies: *small-strip Thin Gap Chambers* (sTGCs) and *Micromegas* (Micro-MESH Gaseous Structure, MMs). Both are able to withstand the higher flux of neutrons and photons expected in future LHC interactions, which will produce counting rates as high as $20\,000\,\text{s}^{-1}\text{cm}^{-2}$ in the inner part of the NSW. These new technologies will greatly improve the ATLAS muon trigger capabilities, allowing for refined event selections.

Additional improvements to the ATLAS muon detector system include 16 new chambers featuring *Small Monitored Drift Tubes* (sMDTs) and *Resistive Plate Chambers* (RPCs) to be installed in the barrel of the experiment, thus improving the overall trigger coverage of the detector. Further, the ATLAS Liquid Argon (LAr) calorimeter will be provided with new front-end electronics and optical-fibre cabling, allowing jets to be better differentiated from electrons and photons and thus greatly improve the resolution of the detector at trigger level.

The first test beams will circulate in the LHC at the end of September 2021, four months after the date planned before the COVID-19 crisis. Provided that the New Small Wheels installation is completed in LS2, Run 3 is scheduled to officially start in March 2022 and has been extended until the end of 2024. The center of mass energy of the collisions is designed to be 14 TeV, the luminosity will reach the value of 2.5 times the nominal one (which is $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$) and it is expected to collect up to $300\,\text{fb}^{-1}$ of data.

2.2 ATLAS: A Toroidal Lhc ApparatuS

ATLAS is one of the four major experiments at the Large Hadron Collider at CERN. It is a general-purpose particle physics experiment run by an international collaboration and it is designed to exploit the full discovery potential and the huge range of physics opportunities that LHC provides. It is designed to take advantage of the unprecedented energy available at the LHC and observe phenomena that involve highly massive particles which were not observable using earlier lower-energy accelerators. ATLAS, together with CMS, was one of the two LHC experiments involved in the discovery of the Higgs boson in July 2012 [28, 40] (Figure 2.6). It was also designed to search for evidence of theories of particle physics beyond the Standard Model.

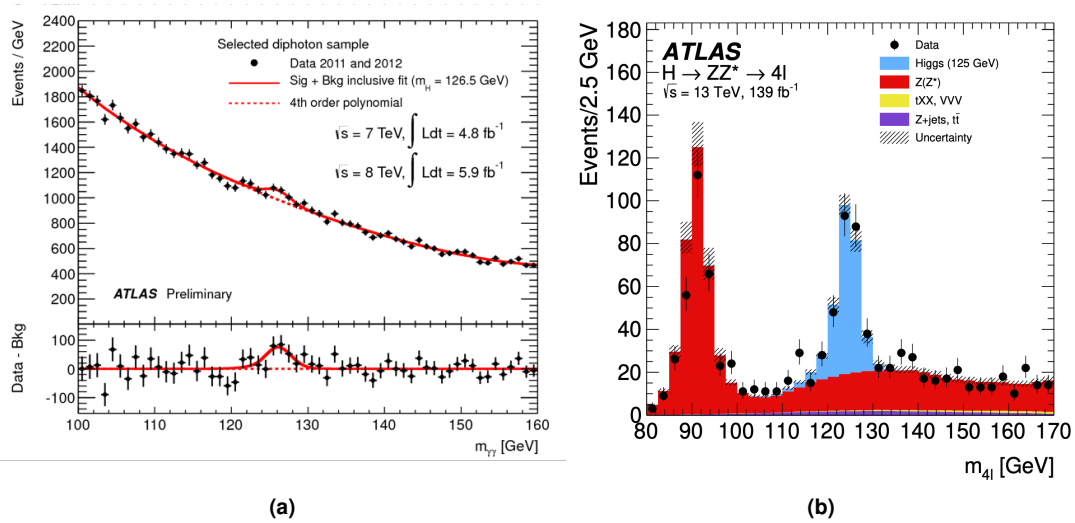


Figure 2.6: (a): The invariant mass from pairs of photons selected in the Higgs to $\gamma\gamma$ analysis, as shown at the seminar at CERN on 4 July 2012. The excess of events over the background prediction around 125 GeV is consistent with predictions for the Standard Model Higgs boson. (Image: ATLAS Collaboration/CERN). (b): A recent [25] distribution of candidate Higgs events from the H to ZZ to 4 leptons analysis using 13 TeV data from the LHC.

Being the largest volume detector ever constructed for a particle collider, ATLAS has the dimensions of a cylinder, 46 m long, 25 m in diameter. The ATLAS detector weight is about 7000 tonnes, similar to the weight of the Eiffel Tower. It sits in a cavern situated 100 m below ground near the main CERN site, which is close to the village of Meyrin in Switzerland.

A schematic representation of the ATLAS detector is shown in Figure 2.5a. It consists of a cylindrical geometry section around the beam axis, called *barrel*, 42 m long with a radius of about 11 m. The barrel is closed at the ends by two *end-caps*, designed to optimize the detection in the forward region. The system almost covers the entire solid angle around the *interaction point* (IP) by means of several concentric sub-detectors of different types. Each of them measures different kinematic features of the particles generated in the proton-proton (*pp*) collisions.

The ATLAS detector consists of the following layers (from the innermost to the outermost):

- an inner tracking system, designed to measure the momentum of charged particles and the position of the interaction vertices;
- a solenoidal super-conductive magnet that provides a uniform magnetic field along the beam axis;
- an electromagnetic calorimeter (ECAL), for the detection and measurement of electromagnetic cascades induced by photons and electrons;
- a hadronic calorimeter (HCAL), for the detection and measurement of hadron showers and the study of jets structure;
- a muon spectrometer, responsible for the tracking and measurement of penetrating muons with a very high precision;
- an air-coded superconducting toroidal magnet, in charge of providing the muon spectrometer with a magnetic field.

The entire detector is named after the peculiar toroidal form of the latter superconducting magnet.

2.2.1 ATLAS Coordinate System

The coordinate system of ATLAS is a right-handed reference frame with the x -axis pointing towards the centre of the LHC tunnel, and the z -axis along the tunnel. The interaction point defines two regions: a "forward" ($z > 0$) and a "downward" region ($z < 0$). The y -axis is slightly tilted with respect to vertical due to the general tilt of the tunnel (Figure 2.8). Due to the symmetry of the detector, a cylindrical coordinates system reference is actually used in ATLAS, where the polar angle θ is measured with respect to the z axis positive direction, while the azimuthal angle $\phi \in [-\pi, \pi]$ lies in the x - y plane.

The θ variable is not invariant under Lorentz boosts along the z axis. In contrast, the *rapidity* variable defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right). \quad (2.4)$$

is a relativistic invariant. E is the energy of the particle considered, while p_L is the longitudinal (i.e. lying on the z axis) component of the particle momentum. In the ultra-relativistic limit, the rapidity y can be replaced by the pseudo-rapidity η :

$$\eta = \frac{1}{2} \ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (2.5)$$

In fact, at extremely high energies $E \approx p_L$. Pseudo-rapidity equals zero for $\theta = 90^\circ$ and asymptotically increases as $\theta \rightarrow 0^\circ$ (or $\theta \rightarrow 180^\circ$), as shown in Figure 2.9.

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

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (2.6)$$

turns out to be useful in physics analysis and trigger efficiency studies and it is widely used.

Inside a hadronic collider such as LHC, the elements that actually collide are particles composed of partons (quarks and gluons). Consequently, the actual interaction energy $\sqrt{s}$ in the center of mass system

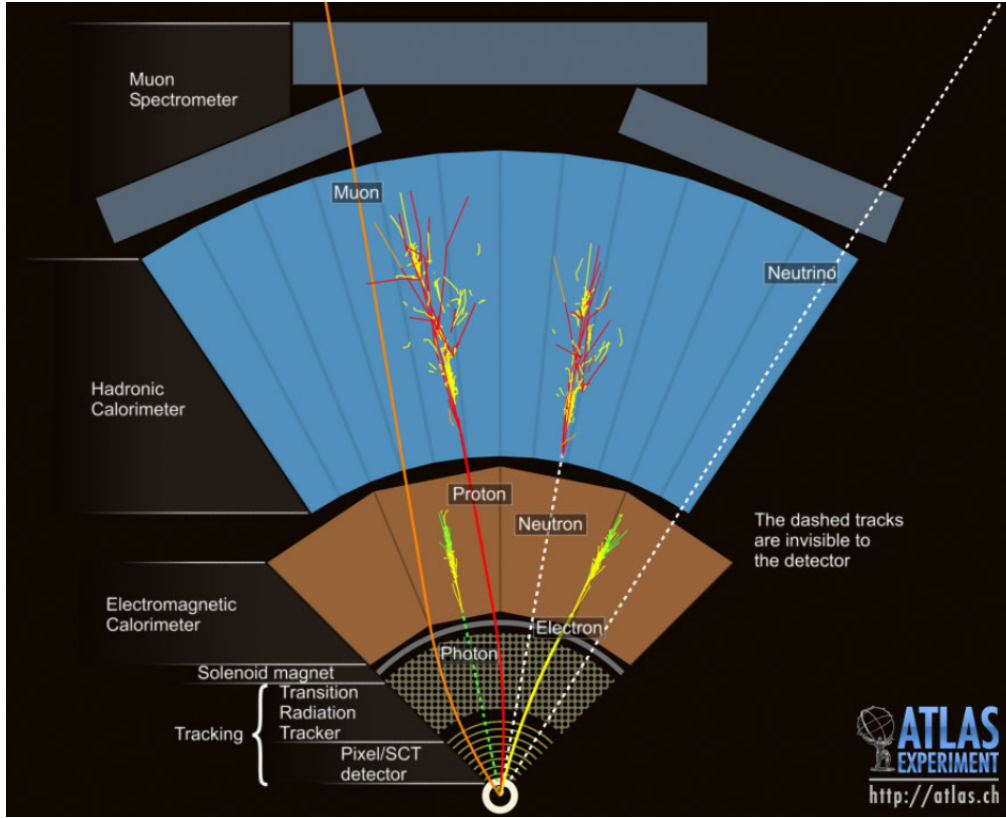


Figure 2.7: A simulated pp collision in the ATLAS detector. The particles produced in the collision are detected by means of their interactions with the components of the ATLAS detectors as they travel through the apparatus.

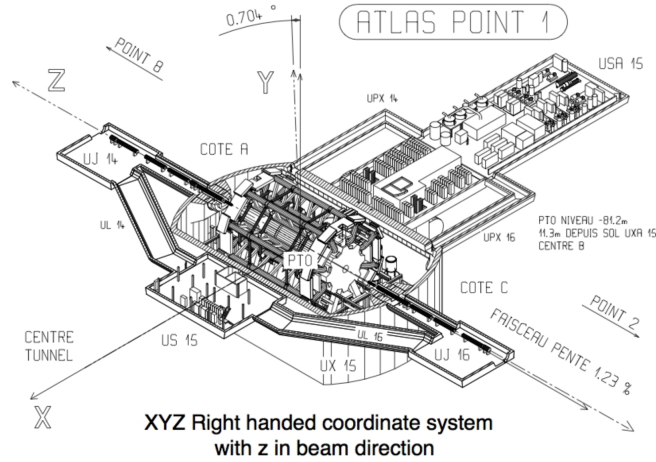


Figure 2.8: Coordinate system used in ATLAS.

is unknown, since it depends on the momentum fraction (x_1 and x_2) carried by the partons that actually participate in the hard scattering process, through the relation (valid when the proton and the parton masses can be neglected)

$$\hat{s} = x_1 \cdot x_2 \cdot s \quad (2.7)$$

where $\sqrt{s}$ is the centre-of-mass energy of the pp system. It is therefore natural to study the kinematics of interactions in the transverse (i.e the x - y) plane. In fact, since the average partons transverse momentum (p_T)

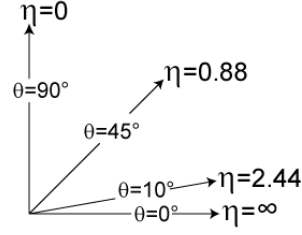


Figure 2.9: Pseudorapidity η for different values of θ angle.

is negligible in the initial state, the energy conservation law allows for the final system p_T being set to zero.

2.2.2 ATLAS magnets

The ATLAS superconducting magnetic system is 26 m long, has a diameter of 20 m and consists of a superconducting central solenoid (CS), which provides the inner detector with a magnetic field, and a system of three superconducting toroids surrounding the CS (Figure 2.10).

The Central Solenoid, with an internal diameter of 2.4 m, covers the central region of the detector and provides an uniform magnetic field of approximately 2 T along the z axis, with a peak of 2.6 T upon the superconductor surface. The solenoidal field bends the tracks of the particles in the transverse plane in order to let the inner tracking system measure their transverse momentum.

The superconducting toroids cover the $|\eta| < 3$ region and have an air-cored structure to minimize the contribution of multiple scattering to the momentum resolution. The toroidal magnetic system consists of two *end-cap* toroids (ECT) and a *barrel* toroid (BT), and provides the muon spectrometer with a magnetic field of 3.9 T for the BT and 4.1 T for the ECTs.

The double magnetic system allows for two independent measurements of the muon momentum: in the inner detector and in the muon spectrometer. This ensures a good muon momentum resolution from few GeV up to the TeV scale.

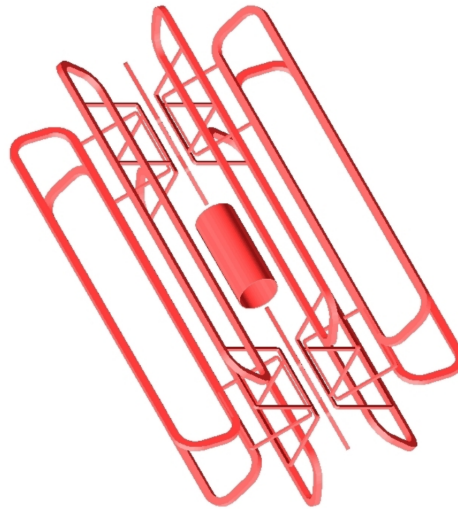


Figure 2.10: The magnetic system of the ATLAS detector: the inner cylinder is the superconducting solenoid, while the external parts are the coils of the toroid.

2.2.3 The Inner Detector

The Inner Detector (ID) is the first part of ATLAS the decay products of the collisions meet along their way, so it is very compact and highly sensitive. The ATLAS physical analyses demand high precision

measurements and require detectors with high granularity because of the enormous track density expected at LHC.

To meet this demand the Inner Detector combines high-resolution detectors in the innermost layer with continuous tracking elements in the outermost layer, the latter being contained within the Central Solenoid.

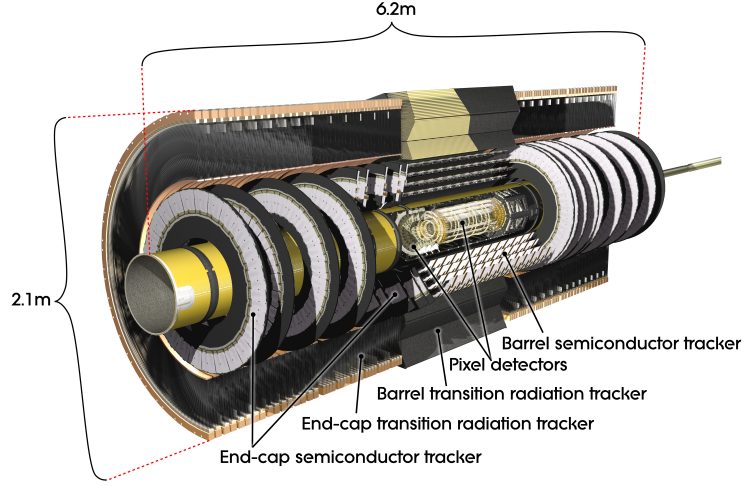


Figure 2.11: The ATLAS Inner Detector structure.

The ID consists of three different sensor systems, all of which are surrounded by a magnetic field parallel to the beam axis (Figure 2.11). The Inner Detector measures the direction, momentum, and charge of the electrically-charged particles produced in each proton-proton collision.

The main components of the Inner Detector are the *Pixel Detector*, the *Semiconductor Tracker* (SCT), and the *Transition Radiation Tracker* (TRT).

- The **Pixel Detector** is composed of three layers of silicon pixels. It guarantees high-precision track measurement, since the spatial resolution on the single hit is $\sim 10 \mu\text{m}$ in the ϕ coordinate and $\sim 115 \mu\text{m}$ along the z coordinate.
- The **SemiConductor Tracker** (SCT) consists of eight layers of silicon strips with a spatial resolution on the single hit of $17 \mu\text{m}$ in ϕ and $580 \mu\text{m}$ along z . The Pixel Detector and the Semiconductor Tracker together provide on average eight high-precision hits per track.
- The **Transition Radiation Tracker** (TRT) is composed of straw tubes chambers. The resolution of such detectors ($\sim 130 \mu\text{m}$ per straw) is lower with respect to semiconductor detectors resolution, 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 collisions alongside with their kinematic features: p_T , η , ϕ and the eventual secondary vertexes due to long-lived particles.

The transverse momentum is measured exploiting the information about the curvature of the track caused by the magnetic field, the latter being provided by the superconducting solenoid described in the previous section. In fact, a particle of charge q and velocity $\vec{v}$, when surrounded by a magnetic field $\vec{B}$, undergoes the Lorentz force

$$\vec{F}_L = q\vec{v} \times \vec{B}. \quad (2.8)$$

Since $\vec{B}$ is along the z axis, $\vec{F}_L$ lies in the x - y plane and causes the trajectory of the particle to bend in the transverse plane. The relation between the transverse momentum and the bend radius r turns out to be

$$r = \frac{p_T}{0.3 \cdot B} \quad (2.9)$$

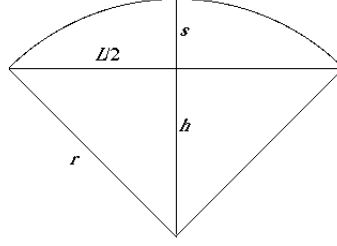


Figure 2.12: Geometric representation of the sagitta.

if r , B and p_T are expressed in m, T, and GeV, respectively. In order to estimate the resolution, the *sagitta method* comes in handy. The sagitta of a track is the maximum distance between the track itself and the straight segment having the same starting and ending points of the track (Figure 2.12). If s is the sagitta, L is the length of the reconstructed track (expressed in meters), N is the number of measured points on the track and ϵ is the resolution on the measurement of the points, the momentum resolution is given by

$$\frac{\Delta p}{p^2} = \frac{\epsilon}{0.3 \cdot B \cdot L^2} \sqrt{\frac{720}{N+4}}. \quad (2.10)$$

As it is made clear from the above formula, it is crucial to have a strong magnetic field, a 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

Calorimeters measure the energy that a particle loses as it passes through the detector. They are also employed to "measure" the missing transverse energy and to perform *particles identification*. A calorimeter is usually designed to entirely stop or "absorb" most of the particles coming from a collision, forcing them to deposit all of their energy within the detector. Calorimeters typically consist of layers of "passive" (or "absorbing") high-density material, for example lead, interleaved with layers of an "active" medium, such as solid lead-glass or liquid argon.

Electromagnetic calorimeters measure the energy of electrons and photons as they interact with matter. Hadronic calorimeters sample the energy of hadrons as they interact with atomic nuclei. Calorimeters can stop most known particles except muons and neutrinos.

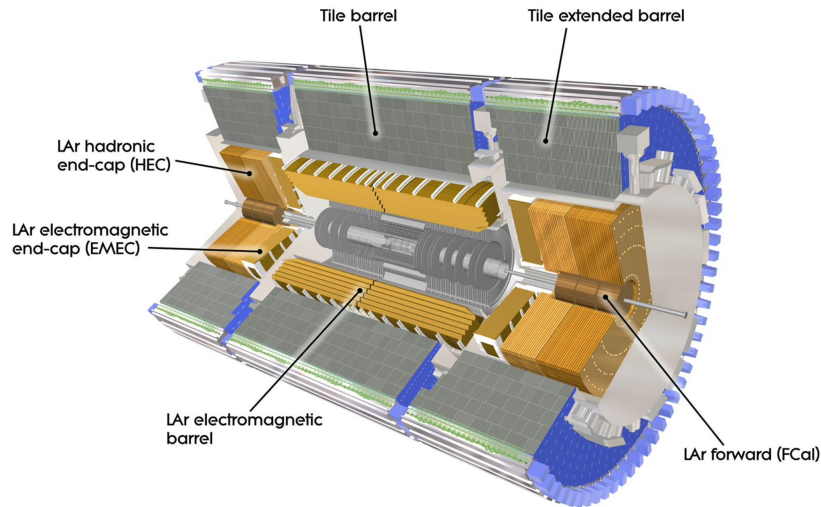


Figure 2.13: The ATLAS Calorimetric system.

The components of the ATLAS calorimetry system are: the *Liquid Argon (LAr) Calorimeter* and the *Tile Hadronic Calorimeter*. A schematic view is given in figure 2.13, where the four parts that constitute the calorimeter can be seen:

- the **electromagnetic calorimeter**, which covers the $|\eta| < 3.2$ pseudo-rapidity range;
- the **cylindrical hadron calorimeter**, which covers the $|\eta| < 1.7$ pseudo-rapidity range;
- two **hadron calorimeters** in the *end-cap* regions, covering the $1.5 < |\eta| < 3.2$ pseudo-rapidity range;
- two **forward calorimeters**, covering the $3.1 < |\eta| < 4.9$ pseudo-rapidity range.

2.2.4.1 Electromagnetic Calorimeter

The *Electromagnetic Calorimeter* is a lead (Pb) and liquid argon (LAr) detector, with an accordion geometry contained in a cylindrical cryostat, which surrounds the inner detector cavity. The accordion structure (made of lead) has a thickness that varies according to η (Figure 2.14), in order to maximise the energy resolution. The active material of the calorimeter is the liquid argon. This structure confers very high acceptance and symmetry in the ϕ coordinate to the calorimeter.

Calorimetric cells are segmented ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) in correspondence of readout electrodes. The longitudinal sampling of the showers is obtained by repeating the cell structure four times along the radial direction. The total thickness amounts to about 25 X_0 (where X_0 is the radiation length) in the *barrel* region and to more than 26 X_0 in the *end-cap* regions.

Overall, the electromagnetic calorimeter provides about 190000 readout channels.

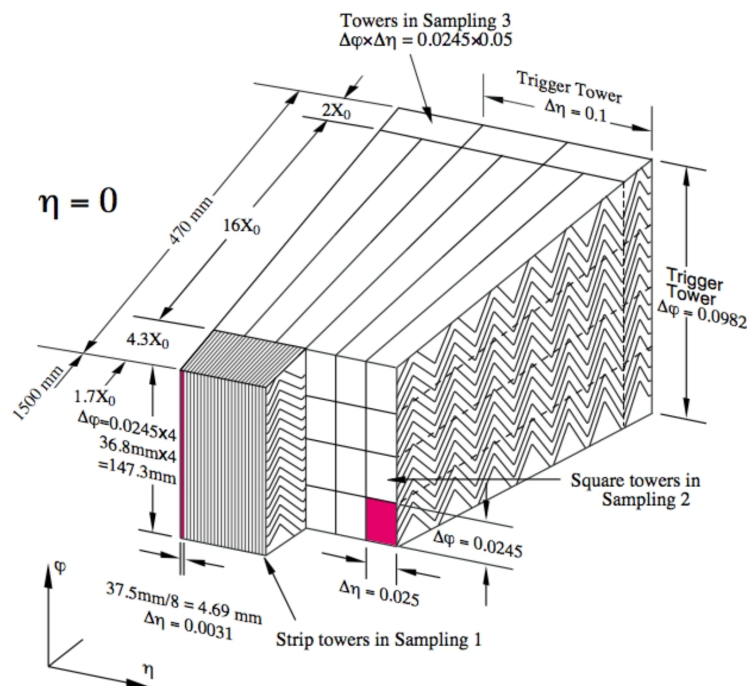


Figure 2.14: The scheme of the accordion structure of the electromagnetic calorimeter with its radial segmentation.

The energy resolution of an electromagnetic calorimeter depends on several factors and can be written as:

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

where a is a *sampling* term which includes the statistical fluctuations, b is a term that takes the noise due to electronics and overlapping signals into account and c is a constant taking account of mechanical effects,

calibration and non-uniform sources which involve systematic errors. For the ATLAS Electromagnetic Calorimeter the energy resolution is equal to $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%$ in the energy range going from 2 GeV up to 5 TeV.

The angular resolution amounts to about $\frac{40 \text{ mrad}}{\sqrt{E[\text{GeV}]}}$, allowing a good measurement in η of the showers direction.

2.2.4.2 Hadronic Calorimeters

The hadronic calorimeters, which cover the $|\eta| < 4.9$ pseudo-rapidity range, use different detecting techniques depending on the region of detecting.

The thickness has been chosen to both provide a good containment of hadronic showers and minimize the number of particles passing through the spectrometer: for $\eta = 0$, the thickness amounts to 8 interaction lengths λ , in order to guarantee a good resolution for high-energy jets and a good measurement of missing transverse energy, E_T^{miss} .

The Hadronic Calorimeters consist in three cylinders which radially extend from an internal radius of 2.28 m up to an external radius of 4.25 m. These cylinders are longitudinally segmented into three layers and divided into 64 modules arranged orthogonally to the beams; hence, the readout channels amount to 10000.

Inside the cylinders, the technique used to measure the energy is the sampling, with an iron absorber that alternates with sparking plates (active material). The structure is periodic along the z axis with sparking plates 3 mm thick and iron plates of total thickness equal to 14 mm in a single period.

The forward calorimeter has a complex structure, since a high radiation level is expected in the region where it is located, about 4.7 m from the interaction point. This is also the reason of its greater propensity to deterioration than the other detectors. It consists of three sections, the first made of copper, the other two of tungsten, and in each of them a metal matrix with channels regularly spaced lengthwise. Again, the active component consists of liquid argon. The total number of readout channels is 3584.

The energy resolution for a hadronic calorimeter can be expressed as

$$\frac{\sigma_E}{E} = \sqrt{\frac{c_{\text{int}}^2 + c_{\text{samp}}^2}{E}} \oplus a. \quad (2.12)$$

where a is a constant which accounts for non-Gaussian fluctuations in the electromagnetic component of the shower, c_{int} allows for the intrinsic fluctuations in the fraction of the initial energy that is transformed into sensible energy, while c_{samp} takes statistical and sampling fluctuations into account.

The energy resolution in the $|\eta| < 3.0$ pseudo-rapidity range is equal to $\frac{50\%}{\sqrt{E[\text{GeV}]}} \oplus 3\%$, while in the range $3.0 < |\eta| < 4.9$ it is equal to $\frac{100\%}{\sqrt{E[\text{GeV}]}} \oplus 10\%$.

2.2.5 The Muon Spectrometer

Muons are particles that usually pass through the Inner Detector and Calorimeter undetected. The *Muon Spectrometer* (MS) identifies and measures the momentum of muons.

The high resolution muon spectrometer is a crucial element of the whole detector, since high momentum muons provide the most clear and simple experimental signatures in many physics searches.

The ATLAS MS has been designed to be independent of the other detectors in order to perform standalone measurements in a wide range of transverse momentum, pseudo-rapidity and azimuthal angle.

The spectrometer occupies both the *barrel* region and the *end-cap* regions, so that a remarkable pseudo-rapidity coverage in the $|\eta| < 3$ range is guaranteed.

The muon spectrometer design was conceived with plans for the detector to meet the following requests:

- a transverse momentum resolution of 1% for the reconstruction of final states having two and four muons and for the identification of muons electrical charge;
- the measurement of the ϕ coordinate with a spatial resolution of 5 mm to 10 mm for the off-line track reconstruction;
- a good trigger discrimination power;

- a good time resolution for the bunch-crossing identification.

The detectors used for the muon tracks reconstruction are of different types (Figure 2.15), depending on the η region, since there exist different rate conditions for different η values.

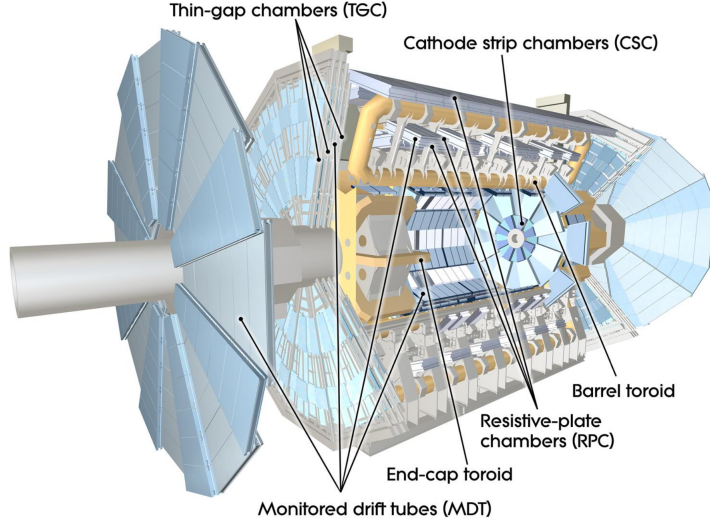


Figure 2.15: The ATLAS Muon Spectrometer, composed of different types of detectors, according to different η region.

Monitored Drift Tubes (MDTs) are employed in the central region ($|\eta| < 2$), while Cathode Strip Chambers (CSC) are exploited for the $2 < |\eta| < 2.7$ range. Additional detectors are installed on the spectrometer for muon detection, with the specific purpose of performing a trigger function: 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 ($\sim$ ns), and a spacial resolution of $\sim$ cm.

In the $r - \phi$ plane the muon system is divided into 16 segments according to the octant symmetry of the toroidal magnet, as it can be seen in Figure 2.16. The structure of the spectrometer system is designed to efficiently exploit the magnets bending power and to cover the entire azimuthal angle ($0 < \phi < 2\pi$).

Three cylinders, which are concentric to the beam axis, are arranged in the barrel chambers. The inner, middle and outer stations cylinder radius is about 5 m, 7.5 m and 10 m, respectively, with the cylinders covering a $|\eta| < 1$ pseudo-rapidity range.

The end-cap chambers, with a trapezoidal shape, cover the $1 < |\eta| < 2.7$ pseudo-rapidity range and are arranged into four disks that are concentric to the beam axis with a distance from the interaction point of 7 m, 10 m and 21 m, respectively.

The Monitored Drift Tubes (MDT) chambers cover a wide pseudo-rapidity range, as illustrated in Figure 2.15. The MDT chambers are composed of aluminium tubes of 30 mm diameter and 400 μ m thickness, with a 50 μ m diameter central wire. The tubes are filled with a mixture of Argon and CO₂ at high pressure (3 bars), and each tube has a spatial resolution of 80 μ m.

Larger values of pseudo-rapidity ($2 < |\eta| < 2.7$) are covered by the Cathode Strip Chambers (CSC). They are multi-wire proportional chambers in which the readout is performed using strips forming a grid on the cathode plane, in both the orthogonal and parallel direction with respect to the wire. The spatial resolution of the CSC is about 60 μ m.

The trigger system covers the $|\eta| < 2.4$ pseudo-rapidity range. The Resistive Plate Chambers, in the barrel, consist of three layers, two of which are located on both sides of the MDT chambers in the middle station and the other one above (below) the MDT in the large (small) sector of the outer station. Along the z axis, the RPC are segmented into 6 or 7 unit of 2.6 m (which equals the MDT segmentation) and they are located in the same mechanical structure to guarantee a good mechanical tolerance and a good alignment with respect to the absolute ATLAS reference system.

In the end-cap region, a vertical arrangement is exploited for the Thin Gap Chambers. A time resolution of 25 ns is required for the trigger chambers to identify the bunch-crossing. The resolution for the second spatial coordinate ϕ is 5 mm to 10 mm.

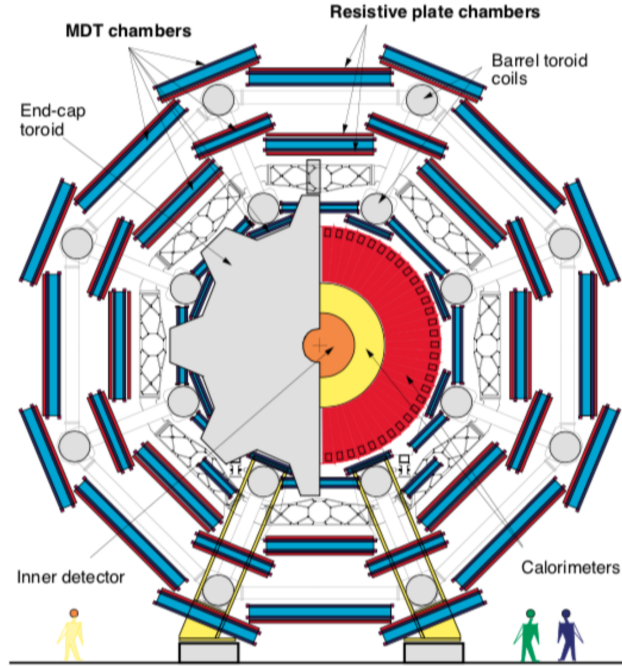


Figure 2.16: The ATLAS Muon Spectrometer section, in the $x - y$ plane.

2.3 The ATLAS Trigger System

ATLAS is designed to observe up to 1.7 billion proton-proton collisions per second, the nominal rate collision into LHC being 40 MHz.

Each event occupies ≈ 1.5 MB, so that a recording rate of ≈ 60 TB per second would be needed to save all information from the ATLAS detector, while the current technology allows for a recording rate of 300 MB/s only. However, not all events contain relevant physical information: the physics processes which are potentially involved with new discoveries are quite rare, so that it is possible to avoid recording a certain amount of events without losing relevant information (Figure 2.17).

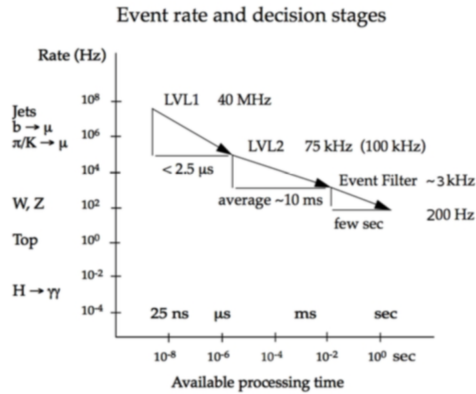


Figure 2.17: Event rate at LHC (as designed) in function of the processing time. As it can be seen, the event rate for rare events of interest (e.g. Higgs in two photons) is covered by characteristic times of trigger processing time.

To reduce the flow of data to manageable levels, ATLAS uses a specialised trigger and data acquisition system.

The ATLAS trigger is designed to rapidly inspect the events detected by the ATLAS detector and

choose whether record or discard the event, after having compared its main features with a set of predefined thresholds contained in the trigger menu.

The trigger system consists in a three-level structure: each level introduces new selection criteria with respect to the previous one and combines information from different detectors, as shown in figure 2.18.

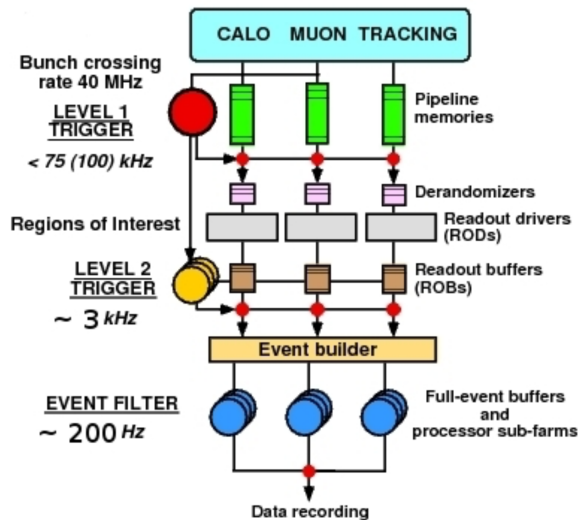


Figure 2.18: Schematic structure of the ATLAS three-level trigger system.

The **first level** of the ATLAS trigger (L1 or LVL1) is completely hardware-based and it uses only information from the calorimetric system and the muon spectrometer. The L1 trigger only looks for high- p_T muons candidates and/or calorimetric objects (electrons, photons and jets), by means of fast and rough measurements performed by ad-hoc detectors in the muon spectrometer (RPC, TGC) and simplified object identifications in the calorimeter. Since it is equipped with the fastest detectors implemented in ATLAS, the first-level trigger is capable of preliminarily reducing the incoming rate from about 40 MHz to less than 100 kHz: L1 trigger is design to decide on one event in $2.5 \mu\text{s}$. The L1 output is a list of so-called *Regions of Interests* (RoI), the latter being regions in the detectors with a specific $\eta - \phi$ range value in which interesting activity has been detected.

The events passing the L1 selection are then retrieved from pipelined storage buffers, before being subjected to the decision of the **second level** of the ATLAS trigger (L2 or LVL2). This trigger is completely software-based. It takes the RoIs provided by the L1 as input and refines the measurements into these regions. The L2 trigger is designed to reduce its incoming rate from ~ 100 kHz down to 3 kHz, taking its decisions in about 10 ms.

Data collected by the precision chambers of the Muon Spectrometers (MDT, CSC) as well as data coming from the Inner Detector are exploited by the L2, while the measurements of the calorimetric objects are refined using higher level algorithms. Also, data collected by several sub-detectors are combined together in order to obtain a better object reconstruction/identification. More specifically, the ID and the MS track information are combined for the muons, while ID and calorimetric information are combined to discriminate between electrons and photons.

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 point, a reconstruction of the whole detector is performed, meaning that the measurement is not restricted to the RoIs anymore. The algorithms exploited are the off-line reconstruction algorithms (adapted to the on-line environment).

EF is capable of taking decisions in about 1 s and the output rate is about 400 Hz. The trigger rates trend as a function of the instantaneous luminosity is displayed in figure 2.19 for the three levels of the HLT. As it can be seen, the trigger rates are stable thanks to changes made in the pre-scales and in the trigger menu: higher thresholds or quality criteria on the trigger objects are required as the luminosity increases.

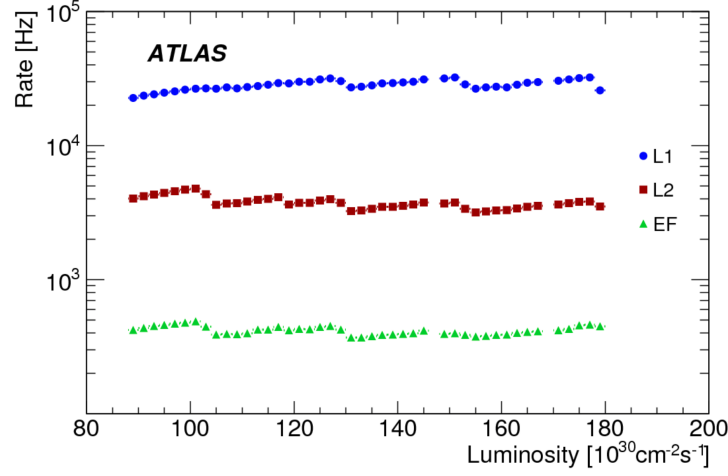


Figure 2.19: Trigger rates as a function of the instantaneous luminosity, for the three levels of the High Level Trigger.

2.3.1 The Electron Trigger

The electron trigger follows the three level ATLAS trigger structure. At first level, the electron trigger makes use of the calorimeters only, hence no distinction between electrons and photons is possible, since they are both identified as "calorimetric objects". L1 exploits a restrained calorimeter granularity. Once a relevant amount of energy is detected, the total energy in a small 2×2 cluster is measured (green area in Figure 2.20), and the isolation with respect to electromagnetic and hadronic activity (yellow and pink areas in the same figure, respectively) is computed. The hadronic activity is due to electrons coming from heavy quark decay. If the E_T , electromagnetic and hadronic isolation parameters fulfil appropriate requirements, the calorimetric object is designated as a good object and its RoI is propagated to the L2.

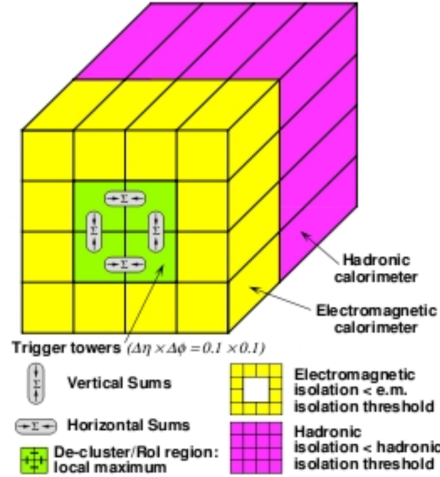


Figure 2.20: Isolation criteria for L1 trigger, for the calorimeter cells: the green area represents the RoIs cluster, the yellow area is the region used for the isolation requirements, and the pink area is the region used for the hadronic isolation.

The L2 trigger basically refines the calorimetric measurement, by making use of the calorimeter full granularity. L2 algorithms are able to study the shape of the energy deposit for the π_0/γ separation using information from the inner tracking system as well. At this level a "calorimetric object" may be identified as an electron if an ID track which is consistent with it is found.

At the end of the chain, the EF further refines the measurements performed at the L2 on the electron candidates, running algorithms that are very similar to the off-line ones and exploiting data from all the

sub-detectors with full granularity.

2.3.2 The Muon Trigger

The L1 muon trigger relies on the temporal and geometric correlation of hits left by a muon on the different layers of RPC detectors, installed in the muon spectrometer (figure 2.21).

When a muon coming from the interaction point comes across the RPC detectors, a hit is recorded by each of them: starting from the hit in the central station (the latter having the function of *pivot* point in the L1 muon identification algorithm), several windows are opened for several p_T thresholds on the RPC1 layer.

If a good hit is found in the RPC1 layer, i.e. if hits in both η and ϕ and in time with the hit on the pivot plane occur, a *low- p_T* muon candidate is identified. 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 muon trigger.

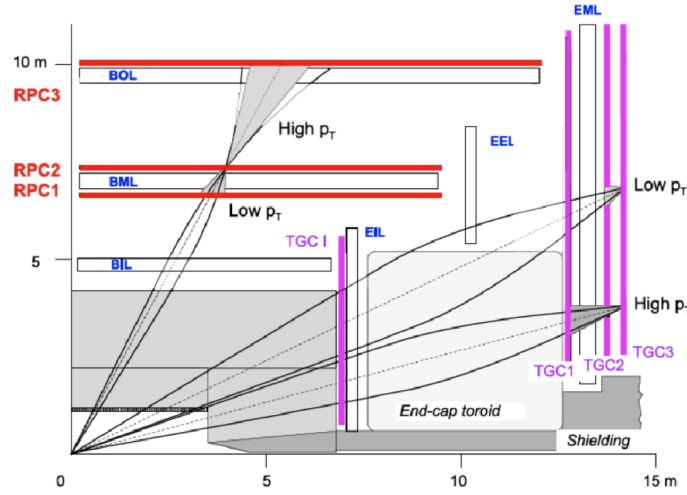


Figure 2.21: The L1 muon trigger algorithm: a muon coming from the interaction point leaves hits on the three layers of RPC detectors. The position of the different hits is correlated with the muon p_T .

At the L2, the muon track is reconstructed for the first time: some algorithms reconstruct the muon track in the Inner Detector (ID) and others in the Muon Spectrometer (MS), separately. There are additional algorithms which combine the two of them, in order to determine the p_T , η and ϕ values associated to the reconstructed track.

It is important to stress that, at this level, the p_T measurement is not performed by means of a fit, but rather using the so-called *look-up tables*. More precisely, the p_T estimation is initially obtained through the equation

$$\frac{1}{s} = A_0 \cdot p_T + A_1 \quad (2.13)$$

where s is the sagitta of the muon track while A_0 and A_1 are two constants taking account of the magnetic field and the energy loss in the calorimeter. A *look-up table* is basically a table whose columns and rows are filled with the (A_0, A_1) pairs found in each cell resulting from the $\eta - \phi$ segmentation of the ATLAS detector. For each muon candidate, at given η , ϕ and s , a quick estimation of the p_T values is therefore obtained. This technique makes it possible to work around the lack of time needed to perform a real fit.

Once the whole 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 the same operations done by the L2 algorithm, but at this level the whole detector with the full granularity can be accessed, hence a real fit of the muon track is performed.

In figure 2.22 the correlation between the muon p_T , reconstructed at different trigger levels, and the off-line reconstruction is displayed: the EF measurements are more accurate and precise than the L2 counterparts.

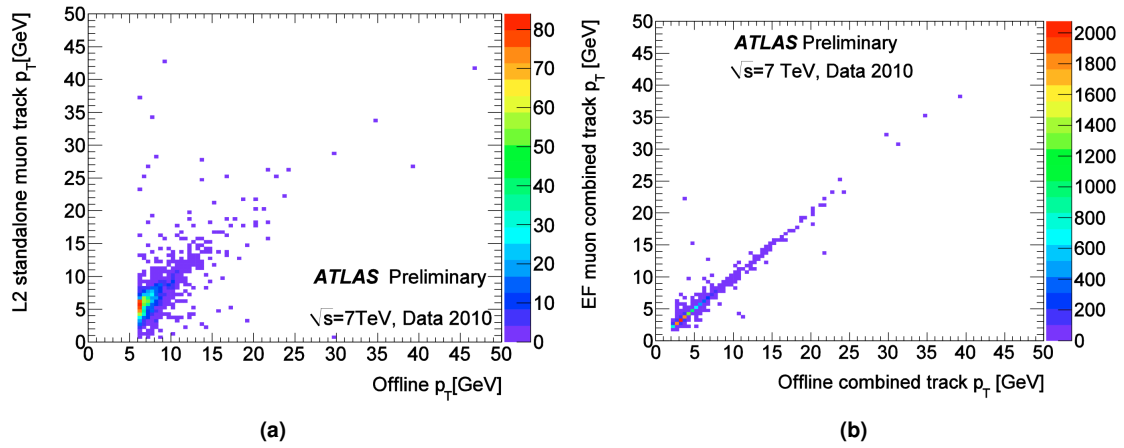


Figure 2.22: Correlation between the muon p_T reconstruction, at L2 in (a) and EF in (b) respectively, and the off-line reconstruction. As it can be seen, the EF measurement is more accurate and precise than the L2 one.

Chapter 3

Physical Object Reconstruction

The outputs of the detectors, generated both by real collisions and simulated ones, are processed by a series of off-line algorithms in order to reconstruct physics objects, whose four-momenta are successively exploited in physics analyses. This chapter focuses on the objects employed in this Thesis, namely electrons, muons and jets.

3.1 Electrons reconstruction

For both electrons and muons, the full reconstruction process proceeds by the following steps:

1. reconstruction;
2. identification;
3. isolation.

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 $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ [17].

An electron from the interaction point loses a significant amount of energy due to bremsstrahlung as it passes through the detector material. The radiated photons may convert into electron-positron pairs which can themselves interact with the detector. These positrons, electrons, and photons are usually emitted in a very collimated beam 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 ID, or can occur downstream of the inner detector, affecting only the shower in the calorimeter. As a result, it is possible to match multiple tracks originating from the same primary electron to the same electromagnetic cluster.

The reconstruction of electron candidates takes advantage of the high-granularity electromagnetic calorimeter and the inner detector, and exploits three fundamental features of electrons signature to form the final electron candidates: 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. Figure 3.1 displays a schematic picture of the elements involved with the reconstruction and identification of electrons.

Electron reconstruction in the precision region of the ATLAS detector ($|\eta| < 2.7$) proceeds by the steps listed below.

1. **Seed-cluster reconstruction:** the $\eta - \phi$ space of the EM calorimeter is divided into a grid of 200×256 elements (*towers*) of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, corresponding to the granularity of the second layer of the EM calorimeter. For each element, the energies (approximately calibrated at the EM scale) collected in the first, second and third calorimeter layers, as well as in the pre-sampler, are summed to form the energy of the tower. Electromagnetic-energy cluster candidates are then seeded from localized energy deposits using a sliding-window algorithm [65] of size 3×5 towers in $\eta - \phi$, whose

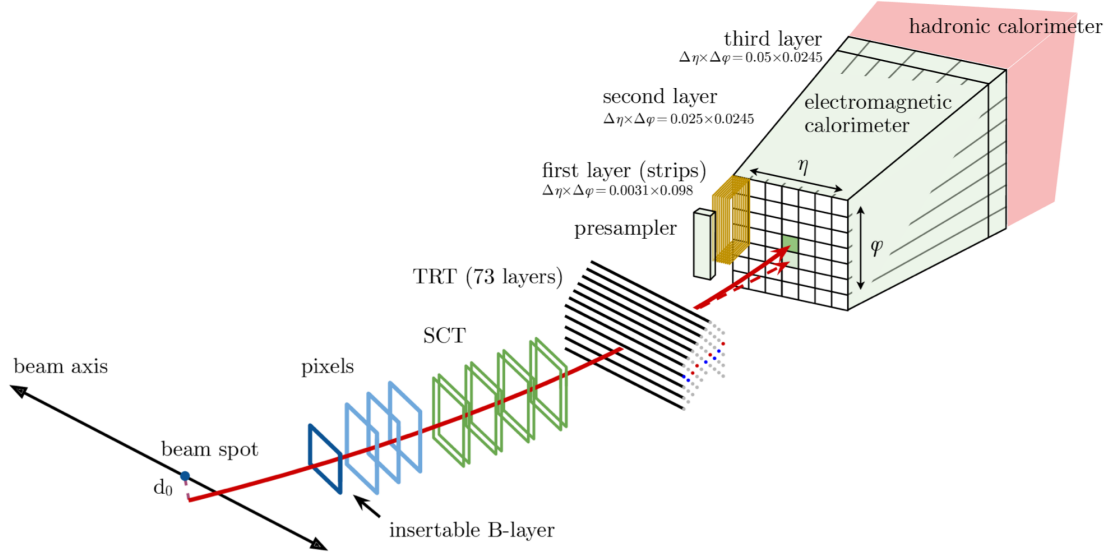


Figure 3.1: A schematic illustration of electrons path through the detector. The red trajectory shows the hypothetical path of an electron, while the dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system.

summed transverse energy exceeds 2.5 GeV. The center of the seed cluster moves in steps of 0.025 in either the η or ϕ direction, searching for localized energy deposits. The seed-cluster reconstruction process is repeated until it has been performed for every element in the calorimeter. If two seed-cluster candidates are found in close proximity, the candidate with the higher transverse energy is retained if its E_T is at least 10% higher than the other candidate. If their E_T values are within 10% of each other, the candidate containing the highest- E_T central tower is kept.

2. **Track reconstruction:** the basic building block for track reconstruction is a "hit" in one of the inner-detector tracking layers. Charged-particle reconstruction in the pixel and SCT detectors begins by assembling clusters from these hits. From these clusters, three-dimensional measurements referred to as space-points are created. The track reconstruction then proceeds in three steps: pattern recognition, ambiguity resolution, and TRT extension [72]. 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 [48]. For tracks which have at least four silicon hits and that are loosely matched to EM clusters in a particular way, a subsequent fitting procedure, using an optimised Gaussian-sum filter (GSF) [62] designed to better account for energy loss of charged particles in material, is applied. The GSF method is based on a generalisation of the Kalman filter [54] and takes the non-linear effects related to bremsstrahlung into account.
3. **Electron-candidate reconstruction:** the matching of the GSF-track candidate to the calorimeter seed cluster candidate and the determination of the final cluster size complete the electron-reconstruction procedure. The track-matching in ϕ is tightened to $-0.10 < q \times \Delta(\phi_{cluster}, \phi_{track}) < 0.05$. If several tracks fulfil the matching criteria, the track considered to be the primary electron track is selected. A further classification is performed using the candidate electron E/p and p_T , the presence of a pixel hit and the secondary-vertex information, mainly for the benefit of keeping a high photon-reconstruction efficiency. Finally, reconstructed clusters are formed around the seed clusters using an extended window of size 3×7 in the barrel region ($|\eta| < 1.37$) or 5×5 in the endcap ($1.52 < |\eta| < 2.47$) by simply expanding the cluster size in ϕ or η , respectively, on either side of the original seed cluster. A method using both elements of the extended-window size is used in the transition region of $1.37 < |\eta| < 1.52$.

Figure 3.2 displays 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.

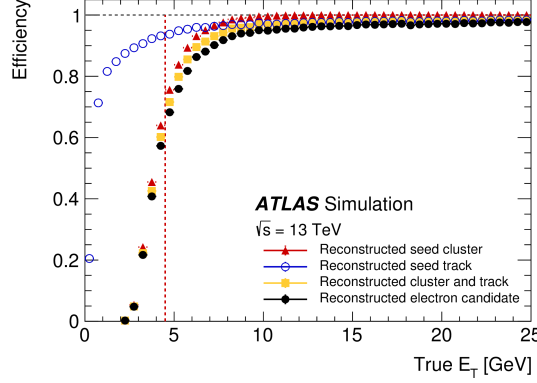


Figure 3.2: Total reconstruction efficiency for simulated electrons as a function of the true (generator) transverse energy E_T for each step of the electron-candidate formation: seed-cluster reconstruction (red triangles), seed-track reconstruction using the Global χ^2 Track Fitter (blue open circles), 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).

3.1.1 Electron identification

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 information from the tracking system or the calorimeter system and quantities that combine both of them (for a complete list of the inputs, see [17]). The method is based on a pseudo-likelihood function that combines all p.d.f. of the inputs, neglecting the correlations

$$L_{S(B)}(\vec{x}) = \prod_{i=1}^N \text{pdf}_{S(B),i}(x_i). \quad (3.1)$$

where S and B subscripts stand for signal and background, respectively. 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 is given by an inverse sigmoid of the likelihood ratio:

$$d'_L = -\tau^{-1} \ln(d_L^{-1} - 1) \quad \text{with} \quad d_L = \frac{L_S}{L_S + L_B}. \quad (3.2)$$

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 3.3a.

A characteristic signature of signal processes at LHC (for example, the decay of heavy resonances such as Higgs, W and Z bosons) is represented by little activity (both in the calorimeter and in the inner detector) in an area of $\Delta\eta \times \Delta\phi$ surrounding the candidate object. However, 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 vicinity of the candidate object, something usually performed by summing the transverse energies of clusters in the calorimeter (E_T^{iso}) or the transverse momenta of tracks (p_T^{iso}) 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 3.3b shows the isolation efficiencies measured in

data and the corresponding data-to-simulation ratios as functions of the electron E_T and η for the operating points given in Table 3.1 and for candidate electrons satisfying Tight identification requirements. The efficiencies that determine the values of the requirements given in Table 3.1 are evaluated in simulation from a $J/\Psi \rightarrow ee$ sample for $E_T < 15$ GeV and from a $Z \rightarrow ee$ for $E_T > 15$ GeV.

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

Table 3.1: 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 $R_{\text{max}} = 0.2$ for track isolation except for the final entry "Fix (Track)" which uses $R_{\text{max}} = 0.4$.

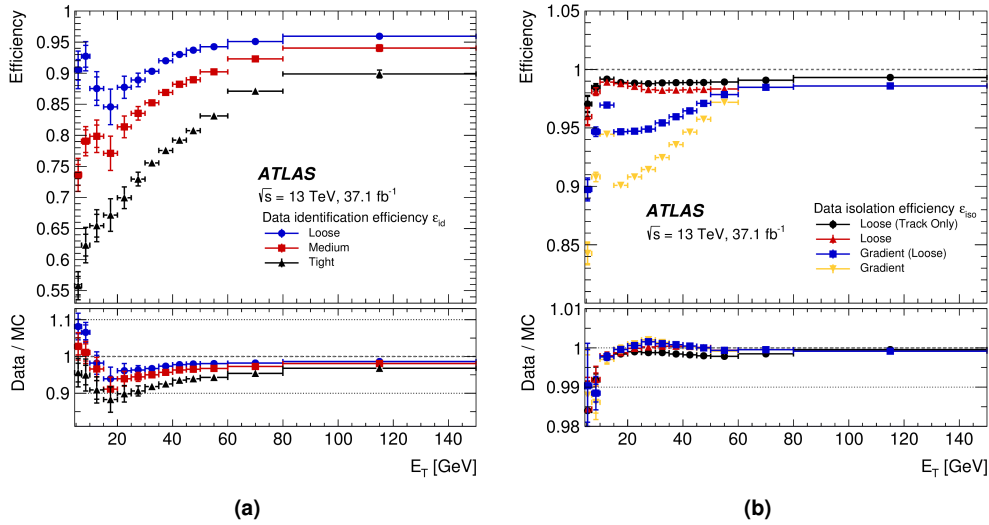


Figure 3.3: (a): Measured LH electron-identification efficiencies in $Z \rightarrow ee$ events for the Loose (blue circle), Medium (red square) and Tight (black triangle) operating points as a function of E_T . The bottom panels shows the data-to-simulation ratios. (b): isolation efficiencies for data (in the upper panels) and the ratio to simulation (lower panels) for the operating points given in Table 3.1 as a function of candidate-electron E_T .

3.2 Muons reconstruction

Muons produced in pp collisions are reconstructed in the ATLAS detector using information from the muon spectrometer (MS), the inner detector (ID) and the calorimeter.

During the 2013–2015 shutdown, the LHC was upgraded to increase the center-of-mass energy from 8 to 13 TeV and the ATLAS detector was equipped with additional muon chambers and a new innermost Pixel layer, providing measurements closer to the interaction point. Moreover, the muon reconstruction software was updated and improved [27].

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. The search algorithm includes a loose requirement on the compatibility of the track with the luminous region.

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.

At least two matching segments are required to build a track, except in the barrel–endcap transition region where a single high-quality segment with η and ϕ information can be used to build a track. The same segment can initially be used to build several track candidates. Later, an overlap removal algorithm selects the best assignment to a single track, or allows for the segment to be shared between two tracks.

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. Hits providing large contributions to the χ^2 are removed and the track fit is repeated. A hit recovery procedure is also performed looking for additional hits consistent with the candidate trajectory. The track candidate is refit if additional hits are found.

Muons pass through calorimeters and reach spectrometer with a large fraction of their initial energies. This feature is used to detect and identify them in muon chambers.

Four muon types are defined, depending on which sub-detectors are used in 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. During the global fit procedure, MS hits may be added to or removed from the track to improve the fit quality. 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 optimised 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.

3.2.1 Muon identification

One of the goals of muon identification is to suppress background as much as possible. Muon candidates originating from in-flight decays of charged hadrons in the ID are often characterized by the presence of a distinctive “kink” topology in the reconstructed track. As a consequence, it is expected that the fit quality of the resulting combined track will be poor and that the momentum measured in the ID and MS may not be compatible.

Several variables offering good discrimination between prompt muons and background muons candidates are studied in simulated $t\bar{t}$ events. Muons from W decays are categorized as *signal muons* while muon candidates from light-hadron decays are categorized as *background*.

For combined muons, the following variables are used to achieve the identification:

- q/p *significance*: it is defined as the absolute value of the difference between the ratios of the charge and momentum of the muons measured in the ID and MS, divided by the the squaring sum of the corresponding uncertainties;
- ρ' : it equals the absolute value of the difference between the transverse momentum measured in the ID and the MS, divided by the p_T of the combined track;
- normalised χ^2 of the combined track fit.

To guarantee a robust momentum measurement, specific requirements on the number of hits in the ID and MS must be met.

Four different "muon" definitions are employed: *Medium*, *Loose*, *Tight*, and *High- p_T* .

- **Medium muons**: in ATLAS, the Medium identification criteria provide the default selection for muons. This selection minimises 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. Specifically, the q/p significance is required to be < 7 .
- **Loose muons**: the Loose identification criteria are designed to maximise the reconstruction efficiency, while providing good-quality muon tracks. They are specifically optimised 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 maximise the purity of muons at the expense of the efficiency. Only CB muons with hits in at least two stations of the MS and satisfying the Medium selection criteria are considered. The normalised 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 maximise the momentum resolution for tracks with transverse momentum above 100 GeV. The selection is optimised 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 suboptimal, 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 3.4 for the four reconstruction criteria.

Muons originating from the decay of heavy particles, such as W , Z , or Higgs bosons, are often produced isolated from other particles, similarly to what happens for electrons. 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, $p_T^{\text{varcone30}}$, is defined as the scalar sum of the transverse momenta of the tracks with $p_T > 1$ GeV in a cone of size $\Delta R = \min(10 \text{ GeV} \times p_T^\mu, 0.3)$ around the muon of transverse

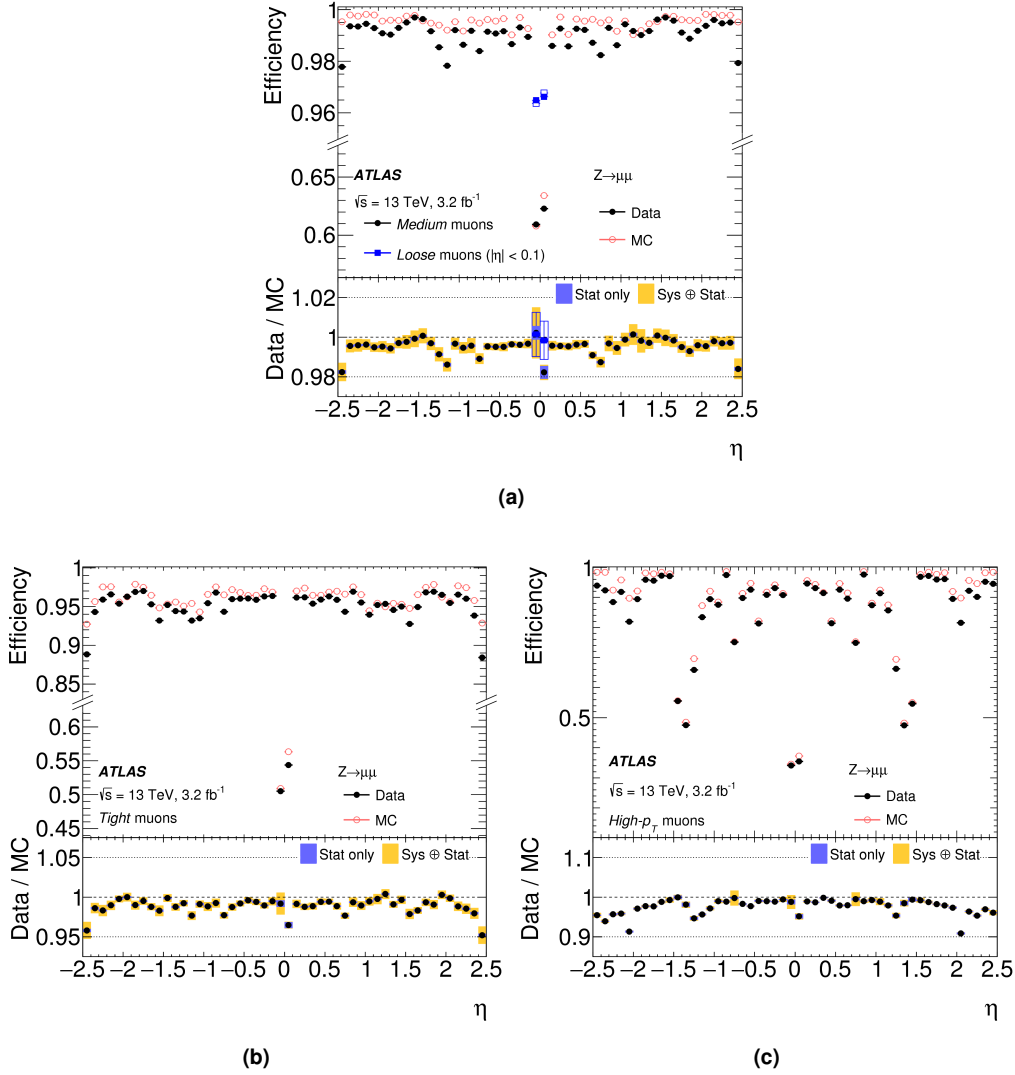


Figure 3.4: Muon reconstruction efficiency as a function of η measured in $Z \rightarrow \mu\mu$ events for muons with $p_T > 10$ GeV shown for the Medium muon selection (a), Tight (b) and High- p_T (c). In addition, the plot (a) also shows the efficiency of the Loose selection (squares) in the region $|\eta| < 0.1$ 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.

momentum p_T^μ , excluding the muon track itself. The calorimeter-based isolation variable, $E_T^{\text{varcone30}}$, is defined as the sum of the transverse energy of topological clusters [19] 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 3.2).

Figure 3.5 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 signal regions 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 muon¹, the opening angle between the two muons, and the ΔR between the probe muon and the closest jet.

¹"Tag" and "probe" refer to the tag and probe method which is used to measure the efficiencies.

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

Table 3.2: 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.

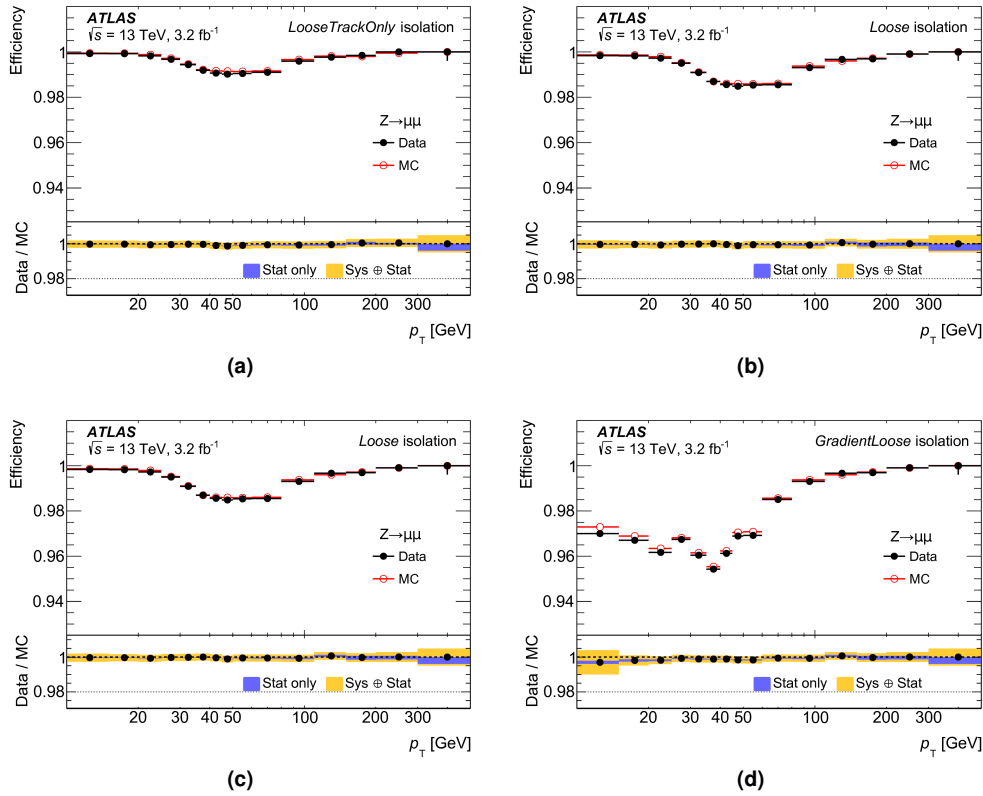


Figure 3.5: Isolation efficiency for the LooseTrackOnly (a), Loose (b), GradientLoose (c) and Fixed-CutLoose (d) muon isolation working points as a function of the muon transverse momentum p_T . The full (empty) markers indicate the efficiency measured in data (MC) samples. The errors shown on the efficiency 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 hadronic colliders such as LHC, the most part of the particles produced in pp collisions consists of hadrons. The latter are formed from quark and gluons, which are usually referred to as *partons*. When two protons collide, only two of such partons are actually involved with the interaction; the rest of the protons constituents participate in generating additional hadrons in the final state. In fact, as already explained in Chapter 1, quarks and gluons cannot exist individually, due to QCD color confinement; as a result, the "spectator" partons (i.e. those not directly involved with the primary interaction) *hadronize* as soon as they

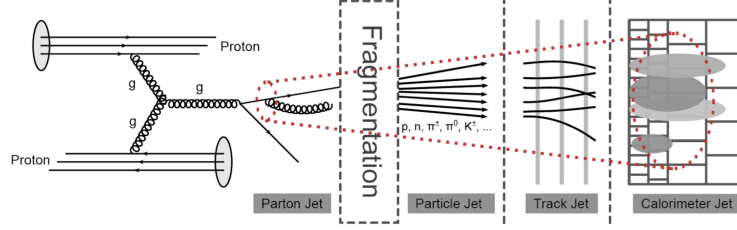


Figure 3.6: Colorful particles, produced at the interaction point, create a bunch of colorless particles due to fragmentation. These particles will produce detector signals in the Inner Detector and Calorimeter, which can be reconstructed as track and calorimeter jets respectively.

are produced. Hadronization is a far from simple process which results in several hadrons being sequentially generated from a single parton; this collimated spray of both charged and neutral hadrons is called a *hadronic jet*.

As they pass through the ATLAS detector, charged hadrons produce ID tracks, while both charged and neutral hadrons deliver energy deposits inside the calorimeters. These trails allow for *track jets* or *calorimeter jets* to be reconstructed, depending on the information used.

The reconstruction process for track and calorimeter jets is the same whether Monte Carlo samples or real data are engaged. Conversely, Monte Carlo samples allow for the definition of two other types of jets:

- *parton jet*: the parton which causes the shower in the calorimeter;
- *particle jet* (or *truth jet*): reconstructed with truth information about particles forming the shower. Only stable particles (with a lifetime greater than 10 ps) are used, i.e. electrons, photons, pions, kaons, protons and neutrons as well as their antiparticles. Neutrinos and muons are not included, since they do not leave any significant signal in the calorimeter.

A schematic view of the different types of jets is shown in Figure 3.6.

Jets reconstruction is built on three steps:

- the definition of calorimeter signals;
- the construction of *topological clusters* starting from calorimeter signals;
- the jet calibration, which allows for jet energy and momentum corrections.

Topological clusters (**topo-clusters**) are three-dimensional and massless calorimeters clusters (i.e. calorimeters adjacent cells with an energy deposit inside) reconstructed by means of a nearest-neighbour algorithm [36]. Cells are added to a topo-cluster according to the ratio of the cell energy to the expected noise in each cell ($|E_{\text{cell}}/\sigma_{\text{noise}}|$) using thresholds that control the growth of each topo-cluster. The resulting energy of the topo-cluster (obtained summing up the energy of all the cells) is defined at the electromagnetic (EM) scale, which is the baseline calorimeter scale that correctly measures energy depositions from electromagnetic showers. A jet produced in the hard-scatter process is expected to originate from the primary vertex, therefore, an event-by-event correction to account for the position of the primary vertex in each event is applied to every topo-cluster [20].

The algorithm used for jets reconstruction is the **anti- k_t** [14] **clustering algorithm** (implemented with the FASTJET 3.2.2 software package [13]). It takes four-vector objects as inputs, such as stable particles defined by MC generators, charged-particle tracks, calorimeter energy depositions, or algorithmic combinations of the latter two, as in the case of the particle-flow reconstruction technique [63].

In the anti- k_t algorithm, two inputs i and j are combined in order to form jets according to a distance parameter defined as

$$d_{ij} = \min \left(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2} \right) \frac{\Delta R_{ij}^2}{R^2} \quad (3.3)$$

where

$$\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2} \quad (3.4)$$

and $k_{T,i}$, y_i and ϕ_i are the transverse momentum, the rapidity and the azimuthal angle of the i -th cluster, respectively, while R is a parameter of the algorithm. The latter calculates $\min(d_i, d_{ij})$. If $\min(d_i, d_{ij}) = d_i$, the input i is said to form a jet and it is removed from the list of inputs. If $\min(d_i, d_{ij}) = d_{ij}$, the inputs i and j are combined into one single input using the sum of four-momentum of each input. The combined input is put into the list of possible inputs, while i and j are removed. The algorithm proceeds until no inputs are left, which means that all inputs in the event will end in a jet.

Calibration is needed since the energy measured in calorimeter clusters does not equal the energy which is actually lost by a particle passing through the detector. Some of the causes are:

- calorimeters non-compensation: the actual energy of a parton that initiates a particle shower is split into an electromagnetic (due to π^0 and η decay into photons) and a hadronic component. Let e and h be the fraction of electromagnetic and hadronic energy, respectively, which is actually detected. If $e/h \sim \mathcal{O}(1)$, a calorimeter is said to be *compensated*. However, in general $e/h > 1$ (as it is for ATLAS), since the hadronic component of the shower is such that some of the energy deposited is "invisible"²;
- energy loss in regions of detector made of passive materials (the so-called "dead" materials of sampling calorimeters);
- leakage, i.e. loss of energy due to particles of the hadronic shower that end up outside the sensitive region of the calorimeter;
- energy deposits of particles belonging to the jet but not included in the reconstructed jet;
- signal loss due to threshold effects, e.g. inefficiencies in calorimeters clustering and jet reconstruction.

The first step in the calibration process is to reconstruct the energy of a jet with respect to electromagnetic scale (EM), which is the basic calorimeter signal scale for the ATLAS calorimeters. A calorimeter signal is first calibrated as the signal comes from an electron. The EM scale is obtained by means of measurements of electrons taken during the test-beam both in the barrel and in the end-cap calorimeters [3, 5, 16, 53] and has been validated using muons signals coming from the test-beams and cosmic-rays. The energy scale of the electromagnetic calorimeters has been corrected using the invariant mass of $Z \rightarrow ee$ events.

This EM scale calibration provides a very good description for energy deposits produced by electrons and photons, but not for deposits from hadronic particles. ATLAS has developed several calibration schemes with different levels of complexity and different sensitivity to systematic effects. Some of them (jet energy scale calibration, jet mass scale calibration, local cell weighting scheme calibration) will be briefly described in the following sections.

Jets can be visualized as "cones" of cascaded particles. The dimension of such cones acts as a discriminant between two types of jets: the *small-radius* jets and the *large-radius* jets. As a matter of fact, default value $R = 0.4$ (small- R jets) is used in jets definition for analyses in ATLAS, but a higher dimension $R = 1.0$ (large- R jets) may be implemented when reconstructing the collimated decay products of boosted weak bosons.

In the following sections, an overview of the jets implemented in this work is given.

3.3.1 Small- R jets

As already mentioned above, small- R jets are reconstructed in ATLAS using the anti- k_t algorithm with a radius $R = 0.4$. The small- R jets used in this work are reconstructed by means of the particle-flow (*PFlow*) technique [22].

²The invisible energy mainly consists of the binding energy of nucleons released in the numerous nuclear reactions, and may represent up to 40% of the total non-em energy, with large event-to-event fluctuations.

3.3.1.1 Reconstruction

Particle flow directly combines measurements from both the tracker and the calorimeter to form the input signals for jet reconstruction, which are intended to approximate individual particles. Specifically, energy deposited in the calorimeter by charged particles is subtracted from the observed topo-clusters and replaced by the momenta of tracks that are matched to those topo-clusters. This improves the accuracy of the charged hadron measurement, while retaining the calorimeter measurements of neutral particle energies. In particular, PFlow jets exhibit improved energy and angular resolution, reconstruction efficiency, and pile-up stability compared to calorimeter jets. After the subtraction, two scalings are applied to account for the difference in response, here defined as the ratio of measured to true particle energy, between topo-clusters at the EM scale and tracks for which the energy scale is closer to the true particle energy. Tracks used in PFlow objects are reconstructed within the full acceptance of the inner detector ($|\eta| < 2.5$), required to have a $p_T > 500$ MeV, and satisfy quality criteria based on the number of hits in the ID sub-detectors. To suppress the effects of pile-up, tracks must also be associated with the primary vertex. Tracks are matched to jets using ghost association [15], a procedure that treats them as four-vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet with which they are clustered [20].

3.3.1.2 Jet energy scale calibration

The jet energy scale (JES) calibration restores the jet energy to that of jets reconstructed at the particle level. In particular, jets are initially calibrated using a sequence of simulation-based corrections. Next, several *in situ* techniques are employed to correct for differences between data and simulation and to measure the resolution of jets.

All JES calibration stages correct the four-momentum, scaling the jet p_T , energy, and mass. The full chain of corrections is illustrated in Figure 3.7 and can be summarised as follows [20]:

- calibrations derived exclusively from MC simulation samples:
 1. pile-up corrections remove the excess energy due to additional pp interactions within the same (in-time) or nearby (out-of-time) bunch crossings. These corrections consist of two components: a correction based on the jet area and transverse momentum density of the event, and a residual correction derived from MC simulation and parameterized as a function of the mean number of interactions per bunch crossing (μ) and the number of reconstructed primary vertices in the event (N_{PV});
 2. the *absolute* JES calibration corrects the jet so that it agrees in energy and direction with truth jets from dijet MC events;
 3. the global sequential calibration (derived from dijet MC events) improves the jet p_T resolution and associated uncertainties by removing the dependence of the reconstructed jet response on observables constructed using information from the tracking, calorimeter, and muon chamber detector systems;
- *in situ* jet calibration: it is applied to correct for remaining differences between data and MC simulation, and it is derived using well-measured reference objects, including photons, Z bosons, and calibrated jets (Figure 3.8).

3.3.1.3 Jet energy resolution

The jet energy resolution (JER) can be parametrized as

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E(\text{GeV})}} \oplus \frac{b}{E(\text{GeV})} \oplus c \quad (3.5)$$

where the $\oplus$ symbol stands for the squaring sum of independent contributions. A stochastic, noise and constant term is represented by the first, second and third addend of the above equation, respectively.

One procedure to measure the JER basically relies on the assumption that dijet events (i.e. events containing only two jets) are such that the p_T values of the two jets are balanced against one other in

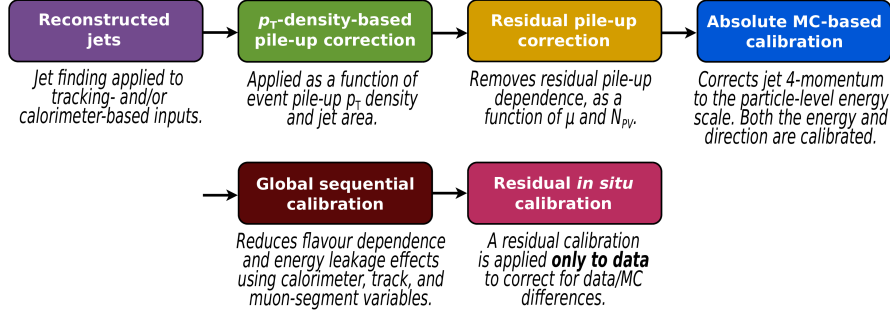


Figure 3.7: Stages of jet energy scale calibrations. Each one is applied to the four-momentum of the jet.

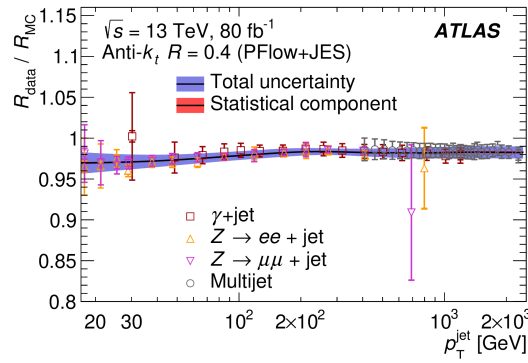


Figure 3.8: Data-to-MC ratio of the PFlow+JES jet response as a function of jet p_T for Z +jet, γ +jet, and multijet in situ calibrations. The inner horizontal ticks in the error bars give the size of the statistical uncertainty while the outer horizontal ticks indicate the total uncertainty (statistical and systematic uncertainties added in quadrature). The final correction and its statistical and total uncertainty bands are also shown, although the statistical uncertainty is too small to be visible in most regions.

the transverse plane due to momentum conservation law. Any deviation from exact balance is due to a combination of experimental resolution, the presence of additional radiation in the event, and biases due to the event selection used in the measurement. In consequence, resolution can be determined by measuring the asymmetry between the p_T values in the same pseudo-rapidity region.

Figure 3.9 shows the JER resolution, alongside with its uncertainty, as a function of p_T for PFlow jets in the central region of the detector, measured using the dijet balance method [20].

3.3.2 Large- R jets

High energy pp collisions can result in the production of massive particles, such as W , Z , H bosons or top quark, with high p_T . When such particles decay into two or three bodies, the typical angular separation between their decay products scales as inversely proportional to the p_T of the initial particle:

$$\Delta R \sim \frac{2m}{p_T}. \quad (3.6)$$

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

In this work, the large- R jets considered are formed using FASTJET implementation of the anti- k_t jet algorithm with distance parameter $R = 1.0$, as already mentioned. Topo-clusters used to reconstruct large- R jets include an additional calibration that is achieved by using the Local Cell Weighting (LCW) scheme [36], so that they are referred to as *LCTopo* jets. The latter are based on three-dimensional clusters of calorimeter cells designed to suppress fluctuations due to noise and pileup. The clusters are calibrated to the

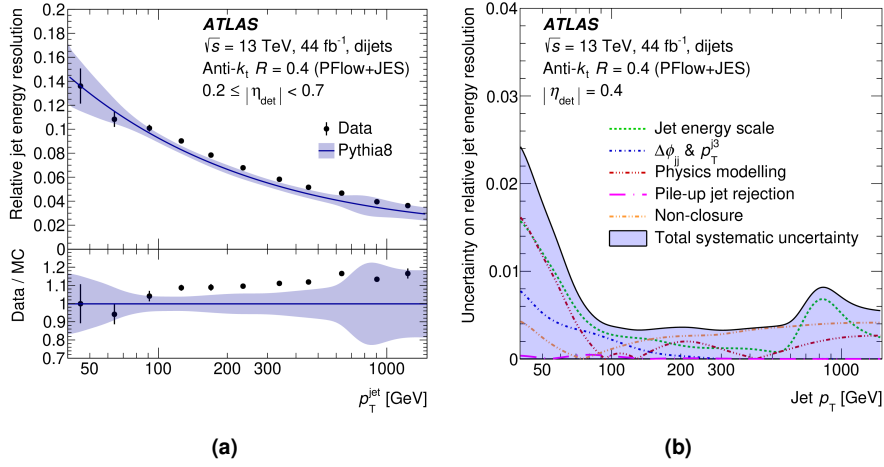


Figure 3.9: (a) Relative jet energy resolution and (b) absolute uncertainty in the relative resolution as a function of p_T for PFlow jets in the central region of the detector, measured using the dijet balance method. The resolution in data is shown in black points with error bars indicating statistical uncertainties; the resolution in detector-level simulated events is shown by the blue curve with total systematic uncertainty given by the blue band.

hadronic scale to take the non-compensation of the calorimeter, out-of-cluster energy, and energy falling in the dead material within the detector into account. Finally, the angular coordinates (η and ϕ) of the cluster are corrected for the selected primary vertex.

In order to suppress contribution from pile-up, the large- R jets undergo a trimming procedure [64], which re-clusters the original constituents of the jet using the k_t algorithm with a smaller radius parameter, R_{subjet} , and therefore produces a collection of sub-jets. The latter are discarded if they carry less than a specific fraction (f_{cut}) of the original jet p_T . The trimming parameters used are $R_{\text{subjet}} = 0.2$ and $f_{\text{cut}} = 5\%$. The remaining sub-jets are responsible for the large- R jet four-momentum.

The different stages of the large- R jet calibration procedure are illustrated in Figure 3.10. 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 [21]. Finally, the jets are calibrated *in situ* (Figure 3.11) using response measurements in data.

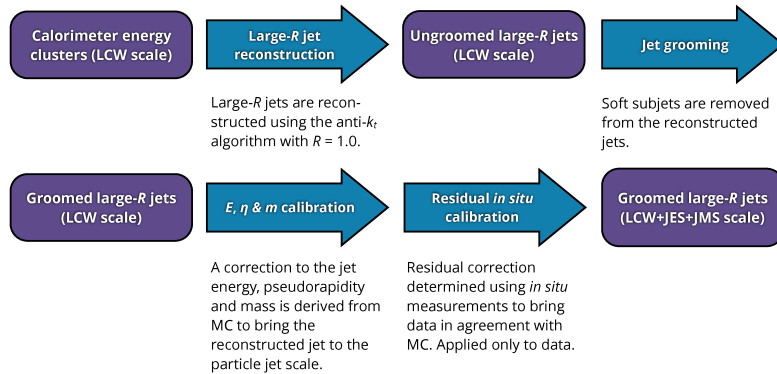


Figure 3.10: Overview of the large- R jet reconstruction and calibration procedure. The calorimeter energy clusters from which jets are reconstructed have already been adjusted to point at the event primary hard-scatter vertex.

The calibration is done taken into account the response of the detector in the measure of the energy (JES, Figure 3.12a) and of the mass (JMS, Figure 3.12b) of the 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.

The large- R jet mass is the result of a combination of calorimeter and tracking information [63]. In fact,

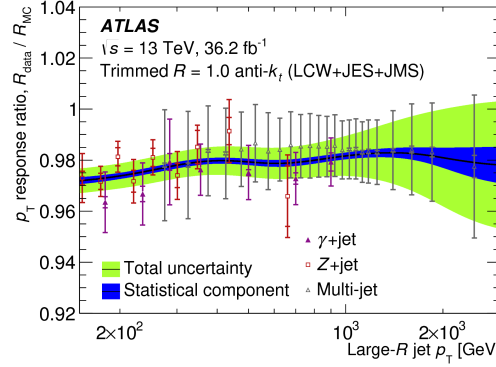


Figure 3.11: Data-to-simulation ratio of the average jet p_T response as a function of large- R jet p_T . The combined result (band) is based on three in situ techniques: the Z +jet balance method (open squares), γ +jet balance method (closed triangles), and the multijet balance (open triangles). The errors represent the statistical (inner error bars) and the total uncertainty (statistical and systematic uncertainties added in quadrature, outer error bars).

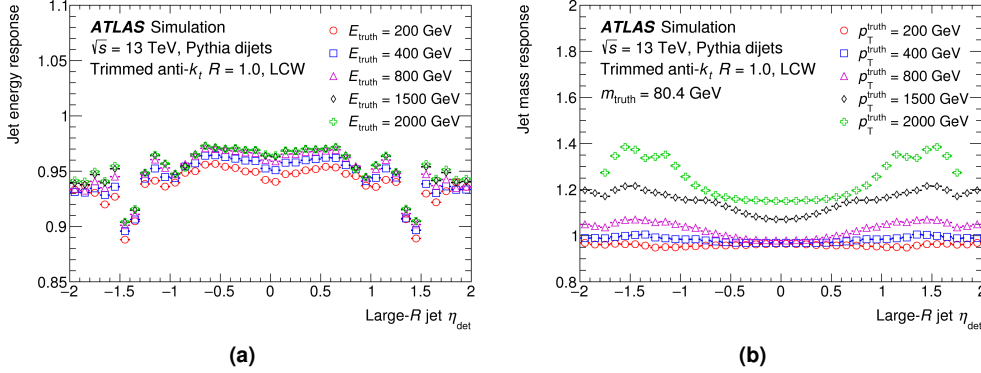


Figure 3.12: Jet energy response (a) and jet mass response (b) as a function of η of the jet.

the mass of a jet is typically defined as the invariant mass of the jet constituents, which are assumed to be massless:

$$m^{\text{calo}} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2} \quad (3.7)$$

where E_i and $\vec{p}_i$ are the energy and the momentum of the calorimeter-cell cluster constituents. The trimmed mass is calculated from the topoclusters that survive the trimming process. However, when the decay products of highly boosted objects are reconstructed in a single topocluster, a better jet mass resolution is obtained by using tracks information, and the resulting mass is referred to as "track assisted mass", defined as

$$m^{\text{TA}} = \frac{p_T^{\text{calo}}}{p_T^{\text{track}}} \times m^{\text{track}} \quad (3.8)$$

where p_T^{calo} is the transverse momentum of a large-radius calorimeter jet, p_T^{track} is the transverse momentum of the four-vector sum of tracks associated to the large-radius calorimeter jet, and m^{track} is the invariant mass of this four-vector sum (the track mass is set to m_π).

Since both trimmed and track assisted mass definitions provide a better mass resolution in different regions of the jet p_T , a weighted linear combination of the two, the so-called combined mass

$$m^{\text{comb}} = w_{\text{calo}} m^{\text{calo}} + w_{\text{TA}} m^{\text{TA}} \quad (3.9)$$

is exploited as jet mass definition which suits all p_T regimes [20].

3.3.2.1 Vector boson tagging

Boson tagging is applied to large- R jets in order to select those consistent with vector boson hadronic decays $V \rightarrow qq$ ($V = Z/W$). Discrimination against background jets inside a mass window including the W/Z mass is based on the variable $D_2^{\beta=1}$, which is defined as a ratio of two-point to three-point energy correlation functions, the latter being built on the energies of and pairwise angular distances between the jet constituents:

$$D_2^{\beta=1} = \frac{e_3^{\beta=1}}{(e_2^{\beta=1})^3} = \frac{\sum_{i < j < k \in J} p_{T,i} p_{T,j} p_{T,k} \Delta R_{ij} \Delta R_{ik} \Delta R_{jk}}{(\sum_{i < j \in J} p_{T,i} p_{T,j} \Delta R_{ij})^3} \quad (3.10)$$

where the sums are taken over the set of subjets in the large- R jet after trimming. This variable is optimised with parameter $\beta = 1$ to distinguish between jets originating from a single parton and those coming from the two-body decay of a heavy particle. In particular, lower $D_2^{\beta=1}$ values are indicative of two-prong large- R jets. More information are given in [66, 67].

The so-called boson tagger is built on p_T -dependent requirements that are simultaneously applied to three variables: m^{comb} , the number of tracks n_{track} associated to the large- R jet and $D_2^{\beta=1}$.

Both the W and Z taggers use a window cut on m^{comb} and one-sided cuts for n_{track} and $D_2^{\beta=1}$. The cut values are optimized in bins of p_T such that the signal efficiency is constant. In this work, two working points are considered, corresponding to a 50% signal efficiency and 80% signal efficiency.

A smooth function is fit to the cut values in p_T bins to produce a smooth tagger with a fixed signal efficiency. The smooth functions used for the fit are

$$\text{fit}_{m^{\text{comb}}} = \sqrt{\left(\frac{A}{x} + B\right)^2 + (Cx + D)^2}, \quad (3.11)$$

$$\text{fit}_{D_2^{\beta=1}} = A + Bx + Cx^2 + Dx^3, \quad (3.12)$$

$$\text{fit}_{n_{\text{track}}} = A + e^{Bx+C}. \quad (3.13)$$

Figure 3.13 show the resulting functions for W tagger, with reference to both the 50% and 80% working points, while Figure 3.14 displays the background rejection (defined as $1/\epsilon_{\text{bkg}}$, with ϵ_{bkg} being the background efficiency) for both the W and Z taggers in reference to both the 50% and 80% working points.

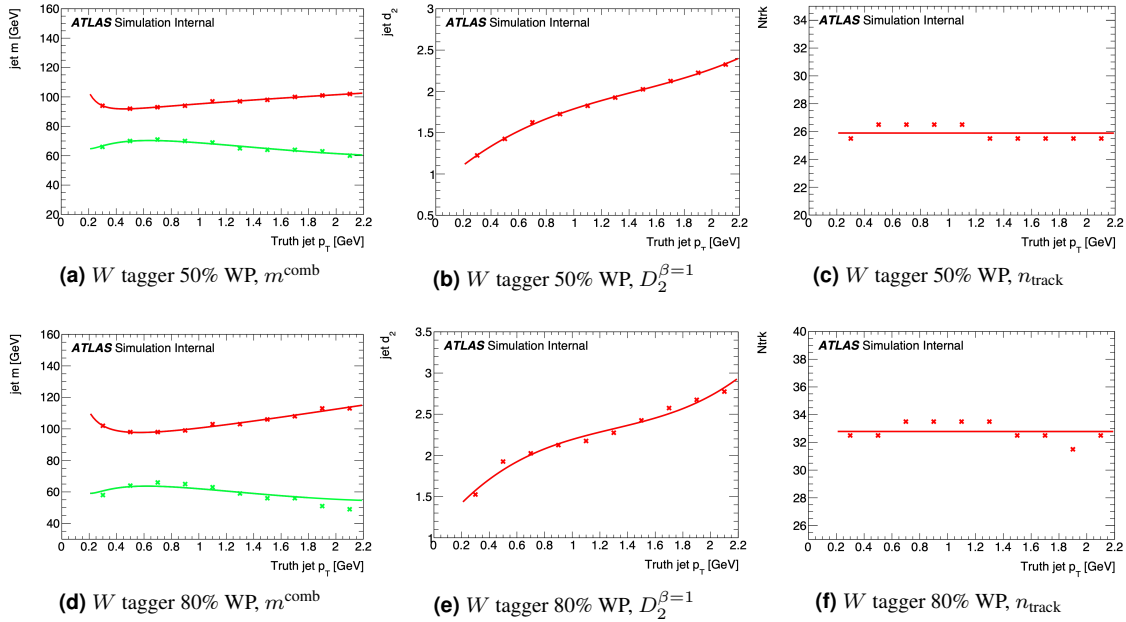


Figure 3.13: W tagger fit functions for m^{comb} , $D_2^{\beta=1}$ and n_{track} , in reference to the 50% working point (first row) and the 80% working point (second row).

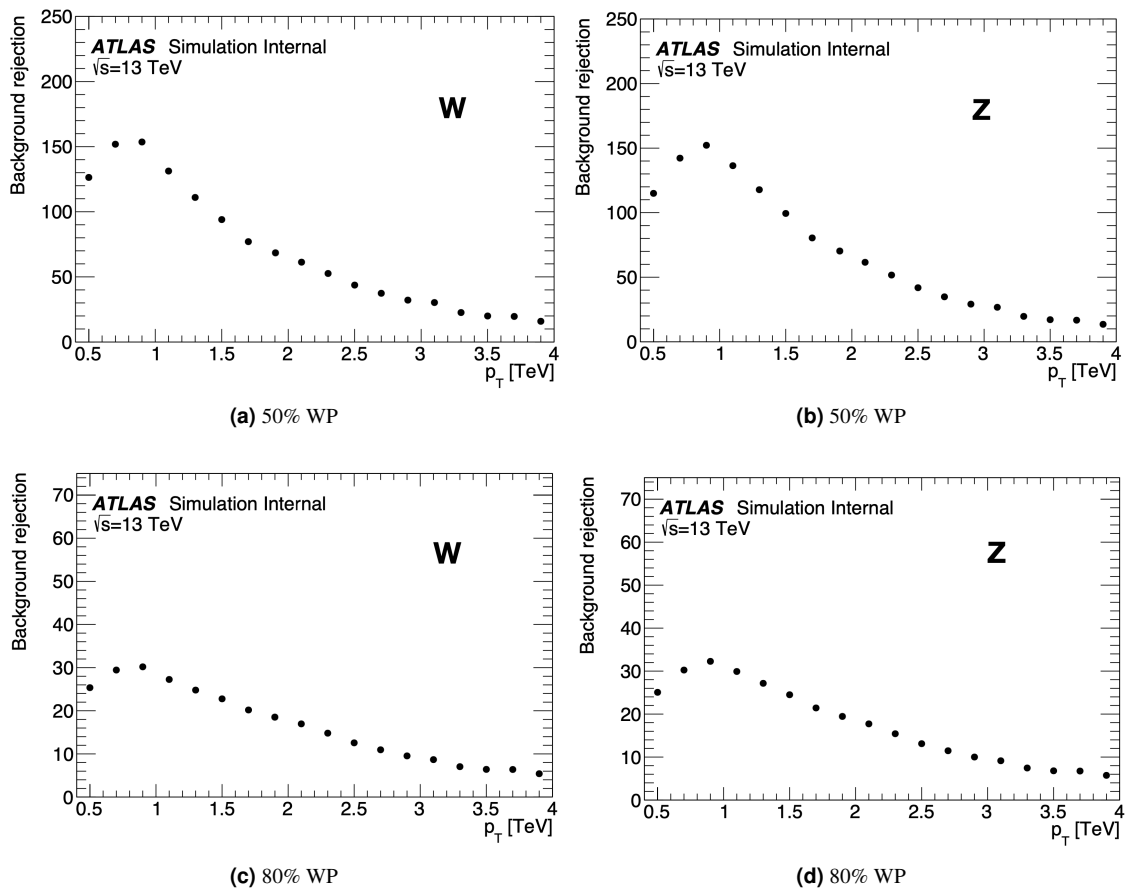


Figure 3.14: Background rejection of W (first column) and Z (second column) taggers, in reference to the 50% (first row) and 80% (second row) working points.

Chapter 4

Search for electroweak diboson production in semileptonic final states

My Thesis work is part of a search for electroweak diboson production in semileptonic final states performed by the ATLAS Collaboration, at the LHC. Indeed, one of the goals of the LHC is to investigate the electroweak symmetry breaking (EWSB), which takes part in the Standard Model in the guise of the Brout-Englert-Higgs mechanism, described in section 1.3.2. A source of electroweak diboson production is vector-boson scattering (VBS), which is involved with interactions between the longitudinal components of the gauge bosons and it is therefore a key tool for probing EWSB. Actually, VBS plays an important role in the exploration of SM physics as well as beyond Standard Model physics: in the framework of a low-energy effective field theory (EFT), VBS cross-section can be modified by anomalous quartic gauge couplings (aQGCs). More precisely, an enhancement of the cross-section at high transverse momentum (p_T) of the vector bosons and at high invariant mass of the diboson system (m_{VV}) is predicted (see section 1.3.5).

At the LHC, VBS is experimentally identified by the presence of a pair of vector bosons (W or Z) and two forward jets with a large pseudo-rapidity separation and a large dijet invariant mass (the whole system is denoted by $VVjj$, where $VV = WW/WW/ZZ$). Vector bosons can decay hadronically ($W \rightarrow qq'$, $Z \rightarrow q\bar{q}$) and leptonically ($W \rightarrow l\nu_l$, $Z \rightarrow l^+l^-$), so that either the

- full-hadronic (both bosons decay hadronically)
- full-leptonic (both bosons decay leptonically)
- semileptonic (one boson decays hadronically, the other boson decays leptonically)

decay channels of the vector boson pair can be picked. While the leptonic channels provide a clean experimental signature, the hadronic ones are characterized by a more favourable branching ratio. The semileptonic channels exhibit the most convenient properties of both the hadronic and leptonic channels. The main challenge of the semileptonic channels is the presence of a large background from W + jets, Z + jets and $t\bar{t}$ events, which make achieving a good VBS signal-to-background ratio very difficult. Conversely, regions of phase space can be found such that these SM backgrounds are largely suppressed with respect to aQGCs signal.

The search for electroweak diboson production in semileptonic final states currently performed by the ATLAS Collaboration is split into three channels, according to the number of charged leptons in the final states:

$V(qq')Z(\nu\nu)$	0-lepton channel
$V(qq')W(l\nu)$	1-lepton channel
$V(qq')Z(ll)$	2-lepton channel

where $V = W, Z$. Each channel is further divided into two different categories:

- "merged", where hadronically decaying vector boson $V(qq')$ is reconstructed as one large-radius jet: $V \rightarrow J$;

- "resolved", where $V(qq')$ is reconstructed as two small-radius jet: $V \rightarrow jj$.

A similar study was already performed by the ATLAS Collaboration, exploiting a dataset recorded with the ATLAS detector at the LHC during the period 2015-2016 and therefore corresponding to an integrated luminosity of 35.5 fb^{-1} [31]. Compared to the previous analysis, the current search is performed using a four-times larger dataset.

My Thesis focuses on the **2-lepton channel analysis**: the final state is $V(qq')Z(ll)$, where ll stands for a lepton-antilepton pair and l may refer to the electronic or the muonic lepton flavour, the tau leptons being excluded from the picture. The current diboson production search is still at a preliminary stage, therefore my contribution has been mainly involved with the optimisation of the event selection. In particular, a machine learning approach involving a Recurrent Neural Network (RNN) has been pursued in order to enhance the signal efficiency while preserving a high signal-to-background ratio in the signal regions. Two different main strategies has been tested: initially, the RNN was trained to recognize VBS-like topology, though training the RNN in order to discriminate signal events against background ones resulted in better performances of the neural network.

Before diving into the details of the analysis, an overview of the Monte Carlo samples and the object reconstruction technique employed in the search is presented.

4.1 Data and Monte Carlo samples

4.1.1 Data samples

The data used in the search were collected with the ATLAS detector in the full LHC Run-2 (2015-2018) pp collisions at a centre-of-mass energy $\sqrt{s} = 13 \text{ TeV}$ and therefore correspond to an integrated luminosity of 139 fb^{-1} . The yearly integrated luminosity is 3.2 fb^{-1} for 2015, 32.9 fb^{-1} for 2016, 44.3 fb^{-1} for 2017, and 58.5 fb^{-1} for 2018.

A mixture of either multiple single-electron or single-muon triggers were employed in order to record the 2-lepton channel events. The single-electron trigger E_T threshold varied from 24 GeV to 300 GeV, while p_T threshold values from 20 GeV to 50 GeV were set for the single-muon trigger. As for the isolation criteria, stricter requirements were applied to low threshold triggers, and vice-versa. On the other hand, no isolation requests were applied to triggers with the highest thresholds. Single-electron triggers were used for the search in the $eeqq$ final state and single-muon triggers were used for the search in the $\mu\mu qq$ final state.

The single-lepton trigger requirements are summarized in Table 4.1.

Data-taking period	$eeqq$ channel	$\mu\mu qq$ channel
2015	HLT_e24_lhmedium_L1EM20 OR HLT_e60_lhmedium OR HLT_e120_lhloose	HLT_mu20_loose_L1MU15 OR HLT_mu50
2016a (run < 302919) ($L < 1.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$)	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e140_lhloose_nod0 HLT_e300_etcut	HLT_mu26_ivarmedium OR HLT_mu50
2016b (run $\geq$ 302919) ($L < 1.7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$)	same as above	same as above
2017	same as above	same as above
2018	same as above	same as above

Table 4.1: List of triggers used in the 2-lepton channel analysis.

4.1.2 Monte Carlo simulation of signal and background processes

Monte Carlo simulations play a key role in particle physics analyses. They are involved with background modelling and estimation as well as the optimization of the event selection and the systematic uncertainties assessment.

The experimental signature of VBS, namely $VVjj$, can be produced by multiple processes. At tree-level, $VVjj$ production has an electroweak (EW) contribution, involving electroweak-interaction vertices only (i.e. amplitudes of $\mathcal{O}(\alpha_{EW}^6)$, where α_{EW} is the weak coupling constant), and a strong (QCD induced) component, including two strong interaction vertices (i.e. amplitudes of $\mathcal{O}(\alpha_{EW}^4 \alpha_S^2)$, where α_S is the strong coupling constant).

EW $VVjj$ production is modelled using MADGRAPH5 2.6 [8]. The EW $VVjj$ samples are generated with the two on-shell V bosons, with one V boson decaying leptonically and the other V boson decaying hadronically. For the W boson, both W^+ and W^- are considered. The purely-electroweak tree-level diagrams considered are:

- EW VBS diagrams, with actual scattering of the two electroweak bosons which occurs via quartic gauge vertices or triple gauge vertices involving the exchange of a Higgs boson or Z/W boson (Figure 4.1);
- EW non-VBS diagrams, without gauge bosons actually scattering (Figure 4.2).

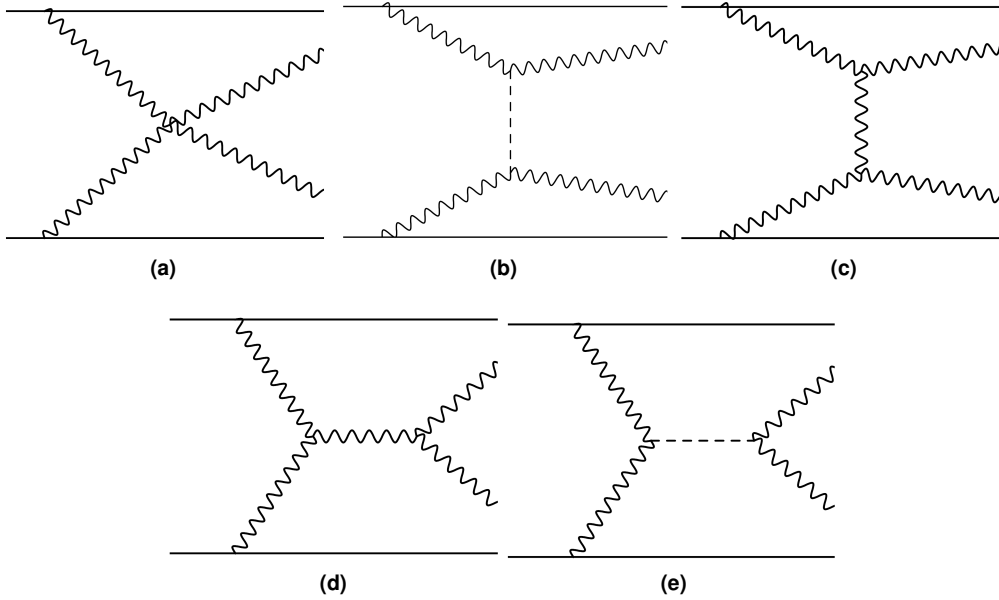


Figure 4.1: Examples of VBS diagrams that contribute to the signal. Not all VBS diagrams contain quartic gauge couplings. The dashed line represents the Higgs boson. The decays of the bosons are not explicitly shown, though they are included in the counting of powers of α_{EW} .

The EW non-VBS component contributes significantly to the total cross-section and it is included in the signal generation, since it cannot be separated by the VBS component in a gauge-invariant way [12]. On the other hand, the background samples account for $VVjj$ tree-level diagrams which contain QCD vertices, such as non-resonant vector-boson pair production (ZZ and ZW , referred to as "diboson" processes, $\approx 1\%$ of the total background, Figures 4.3a, 4.3b) and top quark production (both $t\bar{t}$ pair and single top, $\approx 1\%$ of the total background, Figures 4.3c, 4.3d).

The main background sources for the 2-lepton channel are Z bosons produced in association with jets (Z +jets, $\approx 97\%$ of the total background). Other background sources come from W +jets processes ($< 1\%$ of the total background). Examples of V +jets background tree-level diagrams are displayed in Figure 4.4.

The Z +jets and W +jets processes are simulated with SHERPA 2.2.1 [56] event generator, the matrix elements being calculated for up to two partons at next-to-leading-order (NLO) and up to four partons for leading order (LO). Diboson processes are simulated with SHERPA 2.2.1 as well, for up to one additional parton at NLO and up to three additional partons at LO, while POWHEG-BOX [7] event generator is used to simulate top quark pairs and single-top production processes.

The parton shower, fragmentation, and the underlying events are simulated using Pythia 8.230 [11].

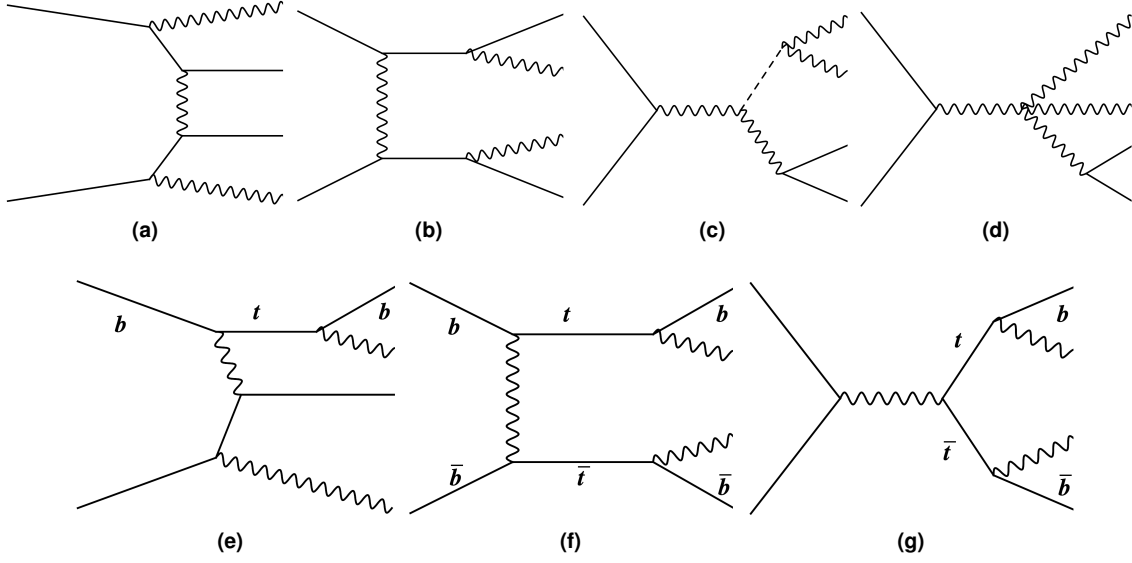


Figure 4.2: Examples of non-VBS ($\mathcal{O}(\alpha_{EW}^6)$) diagrams that contribute to the signal. The decay of the bosons are not explicitly shown, though they are included in the counting of powers of α_{EW} .

Monte Carlo signal and background events are processed with a detailed ATLAS detector simulation [35] based on GEANT4 [4]. Additional overlaid inelastic pp collisions are generated with Pythia 8.186 [73] to model the effects from collisions in the same and neighbouring bunch crossings (pile-up). All simulated events are processed with the same trigger and reconstruction algorithm as the data.

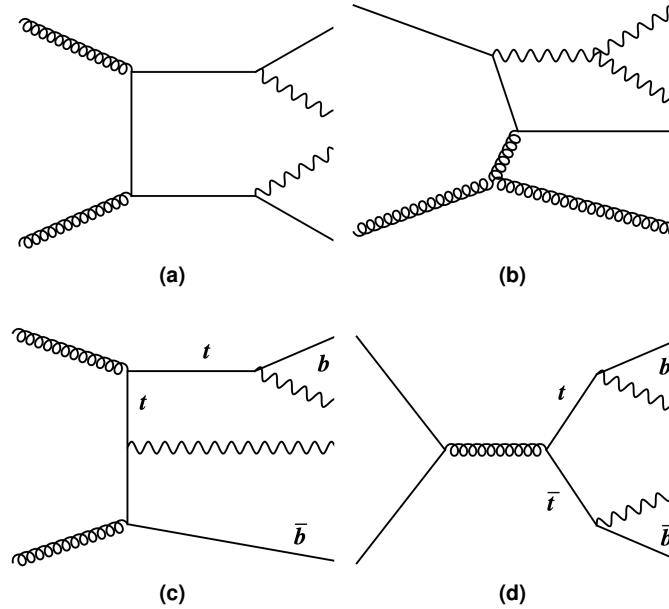
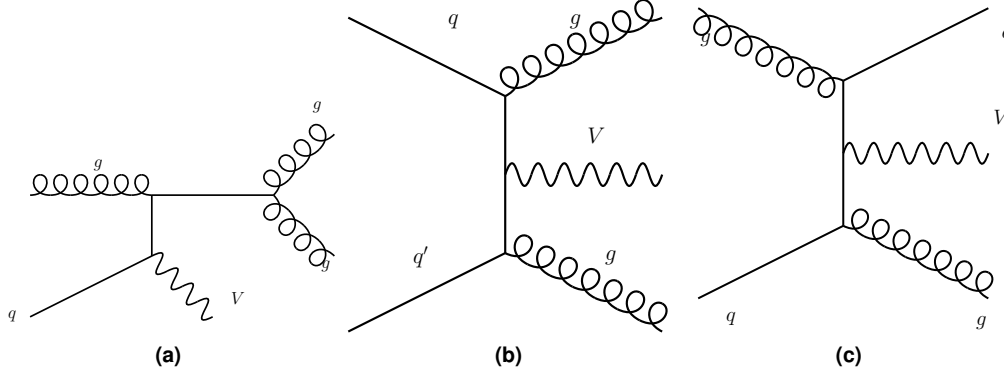


Figure 4.3: Examples of $\mathcal{O}(\alpha_{EW}^4 \alpha_S^2)$ background diagrams. (a) - (c): $V + \text{jets}$ processes; (a), (b): diboson processes; (c), (d): top processes. The decays of the bosons are not explicitly shown, but the counting of powers of α includes the boson decays.

Figure 4.4: Examples of V +jets background tree-level diagrams.

4.2 Object definition

The identification requirements of the physical objects employed in the analysis, namely electrons, muons and jets, are briefly summarized in this section.

4.2.1 Leptons

Electrons are identified as isolated energy clusters in the electromagnetic calorimeter matched to ID tracks. The candidate leptons coming from the $Z \rightarrow ll$ (Z_{lep}) decay are required to be "loose" (section 3.1.1) and to have $p_T > 7$ GeV, $|\eta| < 2.47$.

Muons are reconstructed by a combined fit to the ID and the MS tracks and are required to have $p_T > 7$ GeV and $|\eta| < 2.5$. They must pass identification requirements based on the number of hits in both the ID and the MS. If q is the charge of the muon and $p_{ID}(p_{MS})$ is the momentum of the muon measured in the ID (MS), the significance of the difference $|q/p_{MS} - q/p_{ID}|$ is also involved with the identification requirements definition. Similarly to the electrons, the muons employed in the 2-lepton analysis are classified as "loose" (section 3.2.1).

Leptons from $Z \rightarrow ll$ decays are expected to be isolated from other energy deposits in the detector. Thus, the reconstructed Z_{lep} lepton candidates must satisfy isolation requirements, which are based only on the ID track variable in order to keep high signal efficiency in the high- p_T region, where two lepton candidates are close to each other. Consequently, the FCLoose isolation working point is implemented only at $p_T < 100$ GeV.

As for the associated tracks, lepton must satisfy the conditions $|d_0/\sigma_{d_0}| < 5(3)$ and $|z_0 \times \sin \theta| < 0.5$ mm in the case of electrons (muons), where d_0 is the transverse impact parameter relative to the beam line, σ_{d_0} is its uncertainty, and z_0 is the distance between the longitudinal position of the track along the beam line at the point where d_0 is measured and the longitudinal position of the primary vertex.

A summary of the loose lepton requirements is given in Table 4.2.

	Electrons	Muons
Identification	Loose	
Isolation	FCLoose at $p_T < 100$ GeV and no isolation requirement at $p_T > 100$ GeV	
p_T threshold	7 GeV	
$ \eta $	< 2.47	< 2.5
$ d_0/\sigma_{d_0} $	< 5	< 3
$ z_0 \sin \theta $	< 0.5 mm	

Table 4.2: Selection criteria for leptons used in the 2-lepton analysis.

4.2.2 Small- R jets

Small- R jets (section 3.3.1) are reconstructed from three-dimensional topological clusters of energy deposits in the calorimeter using the anti- k_t algorithm with a radius parameter $R = 0.4$. The small- R jets employed in this analysis combine tracking and calorimeter information with the PFlow technique.

Jets are required to have $p_T > 20$ GeV for $|\eta| < 2.5$ and $p_T > 30$ GeV for $2.5 < |\eta| < 4.5$. A jet vertex tagger [29] is applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$ in order to select only jets associated with the primary vertex while suppressing jets from pile-up interactions. The information exploited by this tagger is about tracks associated with the primary vertex and pileup vertices.

Small- R jets selection criteria are summarized in Table 4.3.

Signal small- R jets	
Input constituent	PFlow
Algorithm	anti- k_t , $R = 0.4$
p_T threshold	20 GeV for $ \eta < 2.5$, 30 GeV for $ \eta \in [2.5, 4.5]$
Pile-up removal	JVT < 0.59 when $ \eta < 2.4$ and $p_T < 60$ GeV

Table 4.3: Selection criteria for small- R jets.

4.2.3 Large- R jets

Large- R jets (section 3.3.2) are employed to reconstruct high- p_T (boosted) massive bosons, which decay into two jets that are very close to each other and can not be reconstructed as two separate small- R jets.

The anti- k_t algorithm is implemented with a radius parameter $R = 1.0$ to reconstruct large- R jets; the inputs to the algorithm are found among topological clusters. Since the Local Cell Weighting scheme is applied for calibration of the large- R jets employed in this analysis, they are referred to as LCTopo jets.

Large- R jets are required to have $p_T > 200$ GeV, $|\eta| < 2.0$ and a combined mass (which is simply referred to as jet mass) m_J greater than 50 GeV.

In order to select large- R jets that are consistent with $V \rightarrow qq$ decays, the so-called boson tagging procedure (section 3.3.2.1) is followed: a p_T -dependent requirement is applied to the jet mass, the number of tracks associated with the jet (n_{track}) and the jet substructure variable $D_2^{\beta=1}$. The latter is optimized to distinguish between jets originating from a single parton from those coming from the two-body decay of a heavy particle, so that lower $D_2^{\beta=1}$ values are indicative of two-prong large- R jets.

The boson tagger algorithm is designed to obtain two fixed signal efficiency working points, 50% and 80%. Table 4.4 summarizes the large- R jets selection.

Signal large- R jet	
Input constituent	LCTopo
Algorithm	anti- k_t , $R = 1.0$
p_T threshold	200 GeV
$ \eta $	< 2.0
Mass threshold	50 GeV
Tagger	3 variables boson tagger

Table 4.4: Selection for large- R jets.

4.2.4 Overlap Removal

After the object selection is completed, an overlap removal procedure is applied in order to discard selected objects which are believed to represent the same particle. The removal criteria are displayed in Table 4.5.

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
small- R jet	electron	$\Delta R < 0.2$
electron	small- R jet	$0.2 < \Delta R < 0.4$
small- R jet	muon	$n_{\text{track}} < 3$ and (ghost-associated or $\Delta R < 0.2$)
muon	small- R jet	$\Delta R < \min(0.4, 0.04 + 10 \text{ GeV/MuPt})$
large- R jet	lepton	$\Delta R < 1.0$

Table 4.5: Overlap removal criteria. The above cuts are applied sequentially. ΔR is calculated using rapidity by default through the formula $\Delta R = \sqrt{(\Delta y)^2 + (\Delta \Phi)^2}$.

4.3 The cut-based approach

In this work, the event selection is described by means of two different approaches. The first one, which is referred to as the cut-based (CB) approach, is summarized in this section in order to be compared with the second approach, which is involved with a RNN (see section 4.4). The latter is the main focus of this Thesis.

In both cases, a set of pre-selections are applied first, including basic quality cuts (event cleaning) and trigger requirements.

Hereinafter, the objects employed in the analysis are considered as ordered according to their p_T value: the "leading" object is the one with the highest p_T among all the objects which are taken into account, the "sub-leading" object has the greatest p_T after the leading one is excluded, and so forth.

4.3.1 $Z \rightarrow ll$ search

The first step in the event selection flow consists of searching for a $Z \rightarrow ll$ candidate. In the 2-lepton channel, the $Z \rightarrow ll$ candidates are selected by requiring two isolated same-flavour leptons satisfying the loose criteria. Only electrons and muon are considered, since τ lepton decays too quickly and it is not reconstructed as clearly as the other leptons.

The leading (subleading) lepton must have $p_T > 28(20)$ GeV, while the dilepton system invariant mass is required to lie in a region around the Z boson mass m_Z :

$$\begin{aligned}
 &83 \text{ GeV} < m_{ee} < 99 \text{ GeV} && \text{in case of electrons} \\
 &(-0.0117 \times p_T^{\mu\mu} + 85.63) \text{ GeV} < m_{\mu\mu} < (0.0185 \times p_T^{\mu\mu} + 94) \text{ GeV} && \text{in case of muons.}
 \end{aligned}$$

The p_T -dependent requirement on $m_{\mu\mu}$ reduces the loss of selection efficiency at high $p_T^{\mu\mu}$ due to the degraded dimuon invariant mass resolution. Since electron are more affected by charge misidentification, opposite charges are required for the muon pairs only.

Events containing a Z_{lep} candidate are referred to as events passing the "dilepton" selection.

4.3.2 VBS Topology: tagging jets

In order to take account of the VBS peculiar topology (specifically, the presence of a dijet system with a large pseudo-rapidity separation and a high invariant mass) a pair of so-called "tagging jets" is required. After selecting the events with a leptonically-decaying Z candidate, the tagging jets are searched: they must be in opposite hemispheres ($\eta_{\text{tag } j_1} \cdot \eta_{\text{tag } j_2} < 0$) and non- b -tagged; if more than such a pair is found, the one with the highest invariant dijet mass is chosen. For the event to be selected, the tagging jets should have $p_T > 30$ GeV to suppress contributions from pile-up and the dijet invariant mass should satisfy the condition $m_{jj}^{\text{tag}} > 400$ GeV. The latter requirement is the cut which the cut-based approach is named after.

As we shall later see, the search for tagging jets is one of the main differences between the CB approach and the RNN approach.

4.3.3 Selection of $V_{\text{had}} \rightarrow J$ candidate (merged selection)

If the tagging jets pair is found in the event, the hadronically-decaying Z/W candidate (V_{had}) is then searched. When V_{had} is boosted, its decay products can be so close that attempting to reconstruct them as two separate small- R jets fails. In order to retrieve signal efficiency, the merged selection is therefore applied as the first step in identifying a V_{had} candidate: if at least one large- R jet with $p_T^J > 200$ GeV and $|\eta_J| < 2$ is found in the event, the boson tagging is applied in order to select the large- R jets consistent with $V \rightarrow qq$ decays: a p_T -dependent requirement is applied to $D_2^{\beta=1}$, n_{track} and m_J , so that m_J always lies in a mass window centred around the expected value of the boson mass.

Two signal regions (SRs) are defined: one for events passing the 50% working point of the boson tagging requirement (High Purity Signal Region) and one for events failing the 50% but passing the 80% working point requirement (Low Purity Signal Region). Large- R jets are required to satisfy either the Z or the W boson tagging. If more than one such large- R jet is found, the one with the greatest p_T is selected as the signal large- R jet.

4.3.4 Selection of $V_{\text{had}} \rightarrow jj$ candidate (resolved selection)

If the event does not pass the merged signal regions, the resolved selection is tried: events are required to have at least two small- R jets with $p_T > 20$ GeV in addition to the two tagging jets. After the latter are selected, the pair among the list of the remaining small- R jets whose dijet invariant mass m_{jj} minimises $\min(|m_{jj} - m_Z|, |m_{jj} - m_W|)$ and whose leading jet satisfies the condition $p_T > 40$ GeV is picked. The mass window defining the resolved signal region is $64 \text{ GeV} < m_{jj} < 106 \text{ GeV}$.

A summary of the event selection for the CB approach is given in Table 4.6.

Selection	2-lepton merged	2-lepton resolved
	exactly two "loose" electrons or muons	
Leptons	$p_T^{l_1} > 28 \text{ GeV}, p_T^{l_2} > 20 \text{ GeV}$ same flavour (ee or $\mu\mu$) opposite charged leptons for $\mu\mu$ events $83 \text{ GeV} < m_{ee} < 99 \text{ GeV}, (-0.0117 \times p_T^{\mu\mu} + 85.63) \text{ GeV} < m_{\mu\mu} < (0.0185 \times p_T^{\mu\mu} + 94) \text{ GeV}$	
Tagging jets	Non- b -tagged $\eta_{\text{tag}j_1} \cdot \eta_{\text{tag}j_2} < 0$, highest m_{jj} $p_T^{\text{tag}j_{1,2}} > 30 \text{ GeV}, m_{jj}^{\text{tag}} > 400 \text{ GeV}$	
Signal jets	1 large- R jet highest p_T 50% (HP) or 80% (LP) WP boson tagging $D_2^{\beta=1}$, n_{track} and $m_{W/Z}$ $p_T^J > 200 \text{ GeV}, \eta_J < 2$ $\Delta R(l_{1,2}, J) > 1.0$	2 small- R jets $\min m_{jj} - m_{Z/W} $ $64 \text{ GeV} < m_{jj} < 106 \text{ GeV}$ $p_T^{j_1} > 40 \text{ GeV}, p_T^{j_2} > 20 \text{ GeV}$

Table 4.6: Summary of the event selection in the 2-lepton resolved and merged analyses.

4.3.5 Control regions

The expected background processes in the signal regions are checked using MC samples. A control region (CR) for the Z +jets processes (as already stated, they represent $\approx 97\%$ of the total background) is defined. The selections are basically the same as the 2-lepton signal region selections, except that the mass window cut for $V \rightarrow J$ or $V \rightarrow jj$ selection is reversed (Table 4.7).

A schematic representation of the various signal and control regions is shown in Figure 4.5, while a summation of the event flow is displayed in Figure 4.6. When an event fails a given selection, it is tested on the next, according to a "recycling" logic. Only upon failing all selections an event is discarded.

Selection	ZCR merged HP	ZCR merged LP	ZCR resolved
Leptons	exactly two "loose" electrons or muons		
Signal jets	1 large- R jet pass50D2 && pass50ntrack && fail80Mass	pass80D2 && pass80ntrack && fail80Mass	2 small- R jets $50 \text{ GeV} < m_{jj} < 64 \text{ GeV}$ or $m_{jj} > 106 \text{ GeV}$

Table 4.7: Summary of the event selection in the Z +jets control regions (ZCR). The other selections are the same as the 2-lepton signal region selections. "pass50D2" and "pass50ntrack" mean the large- R jet should pass the $D_2^{\beta=1}$ and the n_{track} cuts of the 50% working point (WP) of boson tagging, while "fail80Mass" means the large- R jet should fail the mass cut of the 80% working point of boson tagging.

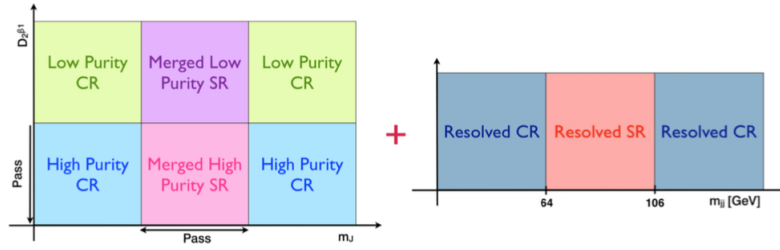


Figure 4.5: Schematic view of the event categorization. There are two merged signal regions (high and low purity), two merged control regions (high and low purity), one resolved signal region and one resolved control region.

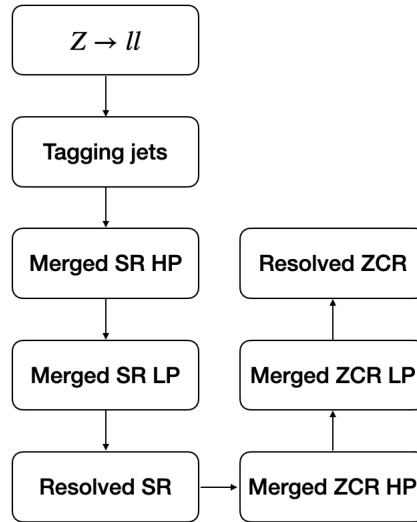


Figure 4.6: Summary of the event categorization in the CB approach.

4.4 The Recurrent Neural Network approach

The implementation of a Recurrent Neural Network (RNN) has been proposed in the current search in order to retrieve a significant amount of signal efficiency in the signal regions, with the goal of making aQGC signal more likely to be observed.

In the cut-based approach, two small- R tagging jets such that the pair invariant mass satisfies the condition $m_{jj}^{\text{tag}} > 400$ GeV are required for the event to be a signal candidate. In the merged regime, at least one large- R jet is required, while in the resolved regime two additional small- R jets are needed to pass the signal region selection. Bearing in mind VBS topology, this is indeed a reasonable choice. Nonetheless, it does not take account of the VBS events such that one or both tagging jets are not reconstructed by the detector due to quality requirements, and therefore result in having less than two small- R jets besides the required large- R jet or the signal small- R jets pair. In addition, if two tagging jet candidates are indeed found, VBS events that fail the m_{jj}^{tag} requirement are discarded as well. As a result for the cut-based approach, the signal efficiency is heavily affected by the search for the tagging jets alongside with the mass cut for the tagging jets pair.

The idea behind using a machine learning-oriented approach is to drop the search for tagging jets and train a neural network to recognize VBS events, exploiting the information about the first small- R jets, up to a maximum of six jets. In fact, VBS signal events differ from background events by the number of small- R jets and their kinematic features (recalling that VBS is characterized by the presence of a high invariant mass and large pseudo-rapidity separation dijet system in addition to the jets coming from the hadronic decay of a boson).

Since each event has a different number of small- R jets, the neural network has to deal with a variable-length input. To address this task, a RNN architecture was chosen.

Given the fact that the input variables are concerned with jets information only, the RNN approach can be easily extended from the 2-lepton to the other channels.

4.4.1 RNN architecture and training

The RNN is built with Keras [6] library using the Theano python library [74] as a back end for mathematical computations. It has already been successfully implemented in the search for heavy diboson resonances in order to separate vector boson fusion (VBF) production processes from gluon-gluon fusion and Drell-Yan processes [32].

The input features of the RNN are the four-momentum variables (p_T, η, ϕ, E) alongside with the number of tracks (n_{track}) of the first small- R jets, up to a maximum of six jets. According to the intrinsic variable length of the input set, the input variables values are fed to the RNN in a recurrent way. The last layers have just one neuron in order for the network to have a single score as output (Figure 4.7). Table 4.8 summarizes the technical setup of the RNN architecture.

Loss function	binary crossentropy
Train:Test:Validation	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 4.8: Technical setup of the RNN architecture.

The distribution of the input variables of the first two jets for ZW signal and Z +jets background events (passing the dilepton selection) are displayed in Figure 4.8.

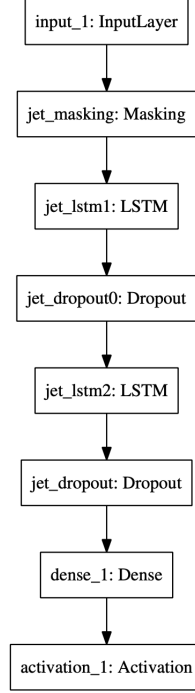


Figure 4.7: Scheme of the RNN architecture.

In a supervised learning framework, the neural network is fed with a training dataset consisting of instances whose desired output (or score) is already known. More specifically, the training dataset contains both signal and background Monte Carlo samples: signal events are labelled as belonging to the "1" category and background events are labelled as belonging to the "0" category. On the basis of the "examples" it is given, the network learns a model to distinguish the "1" category from the "0" category (i.e. it accomplishes a *classification* task), so that it is able to make a prediction about uncategorised events. In reference to the latter, the RNN output (a real value spanning from 0 to 1) is therefore the *probability* for such events to be signal-like. Then, one can arbitrarily choose a threshold value X of the score such that an event is identified as signal-like if the corresponding output is higher than X , and vice-versa.

Usually, the model learnt by the network is tested by means of a test dataset, i.e. a set of labelled instances which are not included in the training dataset. Since the test instances labels are known, one can assess the number of correct classifications made by the network and evaluate the goodness of the model.

Figure 4.9a shows the RNN score comparing the training and the test dataset used. As it can be seen, the distributions for signal and background events are largely separated in both training and test cases, with signal events being concentrated around 1 and background events around 0. The more distant the peaks are, the purer the signal selection is at given X , meaning that the fraction of background events selected as signal ones is restrained. In other words, the more distinct the distributions are, the better the RNN performance is. This concept is encoded by the so-called ROC (Receiving Operating Curve), which is illustrated in Figure 4.9b. By definition, the ROC relates a true positive rate and a false positive rate, respectively defined as:

$$\text{True Positive Rate (TPR)} = \frac{\# \text{ of truth VBS events with RNN score} > X}{\# \text{ of truth VBS events}} \quad (4.1)$$

$$\text{False Positive Rate (FPR)} = \frac{\# \text{ of non-VBS events with RNN score} > X}{\# \text{ of non-VBS events}} \quad (4.2)$$

namely, the fraction of VBS events correctly classified as such and the fraction of non-VBS event classified as VBS events. For each value of X , a pair made of a true positive and a false positive rate (i.e. a point of the ROC) is found. X is usually chosen to maximize the TPR while keeping the FPR small. The ROC of a poor classifier which randomly identifies the events would be a straight line from (0, 0) to (1, 1). If we imagine this straight line to be the diagonal of a square, a good classifier is such that its ROC lies in the lower

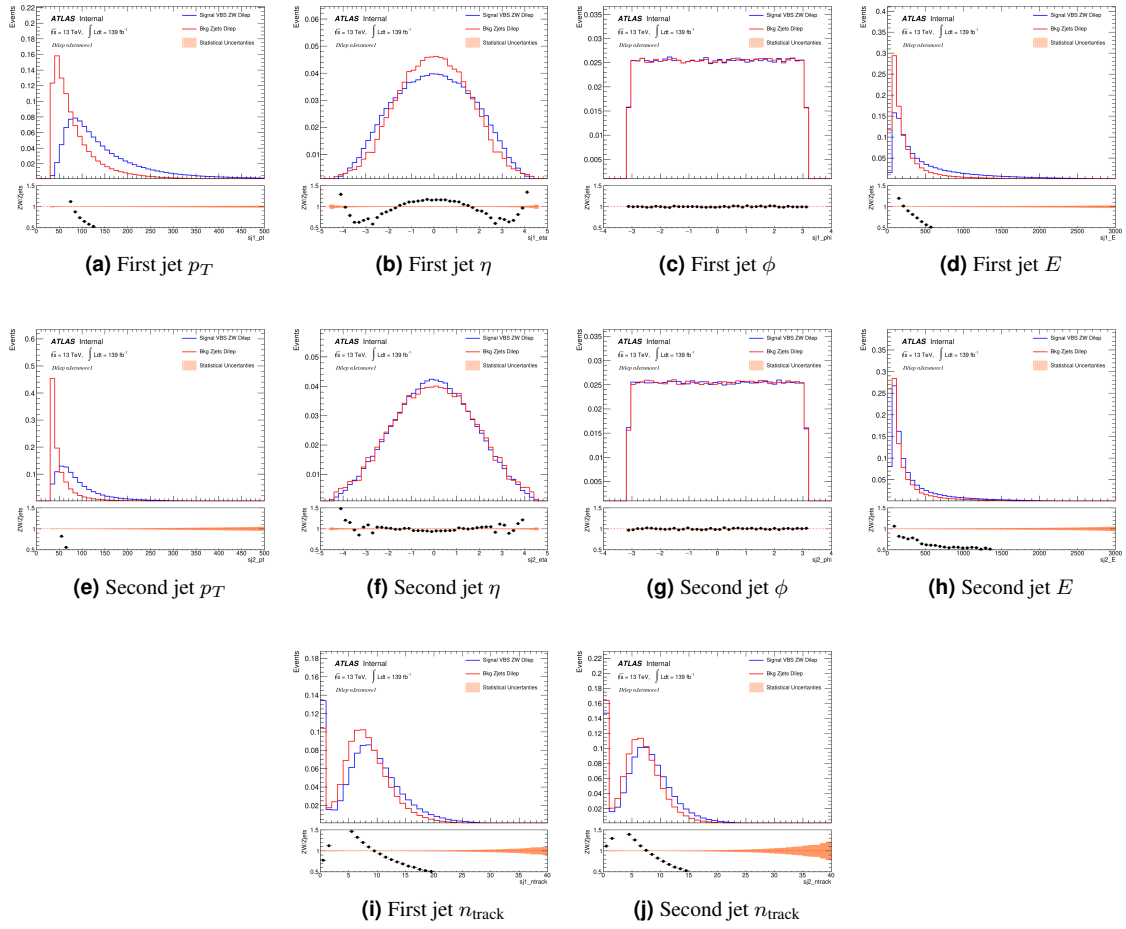


Figure 4.8: Four-momentum variables and n_{track} (normalized) of the first two jets, for ZW signal events (blue curves) and Z+jets events (red curves). The bottom pads display the ratios between the ZW signal histograms and the Z+jets histograms.

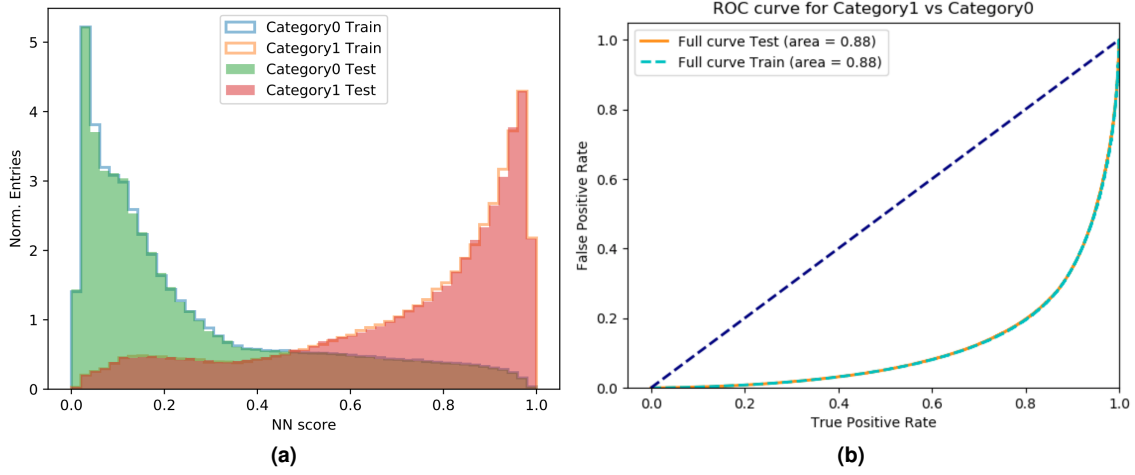


Figure 4.9: (a): RNN score distribution for signal and background events, comparing the training and the test dataset; (b): RNN ROC.

right area of this very square. The higher the separation between signal and background distributions is, the better the RNN is able to discriminate between signal and background events, the closer to the square's lower right vertex the ROC is. To quantitatively describe the latter statement, the AUC (Area Under the Curve) parameter is usually provided as a measure of the network performance (a value as close to 1 as possible is desired).

4.4.2 VBS topology categorization

Initially, the RNN training was performed using events passing the dilepton selection, when at least two small- R jets were found in the events ($N_{\text{jets}} > 1$). VBS topology could be selected by applying a loose cut to the RNN score (RNNScore) instead of searching for the tagging jets pair and then rejecting events such that $m_{jj}^{\text{tag}} < 400$ GeV. Indeed, as one can infer by looking at Figure 4.10, the score of the network peaks around 1 for signal-like events and around 0 for background-like events, so that requiring that RNNScore is greater than 0.3 (for example) results in rejecting a large amount of background events while keeping the most part of signal events.

The event selection and categorization would be the same as the CB approach selection flow, except for making use of the RNN and selecting events such that $\text{RNNScore} > \text{RNNScore}_{\text{cut}}$ (Figure 4.11). The optimal value of $\text{RNNScore}_{\text{cut}}$ was searched in order to maximize both signal efficiency

$$\epsilon_S = \frac{\# \text{ of signal events passing a selection}}{\# \text{ of signal events passing dilepton selection}} \quad (4.3)$$

and the expected asymptotic (or "binned") significance, the latter being estimated as

$$\sigma_{\text{asy}} = \sqrt{\sum_{i=1}^{N_{\text{bin}}} 2 \cdot [(S_i + B_i) \cdot \ln(1 + S_i/B_i) - S_i]} \quad (4.4)$$

where N_{bin} is the number of bins of the histogram chosen to compute the binned significance, while S_i and B_i are the number of signal and background events in the i -th bin, respectively. From a statistical point of view, an excess S of events with respect to the expected background B is said to be *significant* if S is unlikely ascribed to background statistical fluctuations. As regards the asymptotic significance, it makes sense to calculate it on a discriminant variable, i.e. a variable which exhibits different shapes for signal and background events. Sometimes, a less accurate (but easier to calculate) estimation of the significance is provided:

$$\sigma = \frac{S}{\sqrt{B}} \quad (4.5)$$

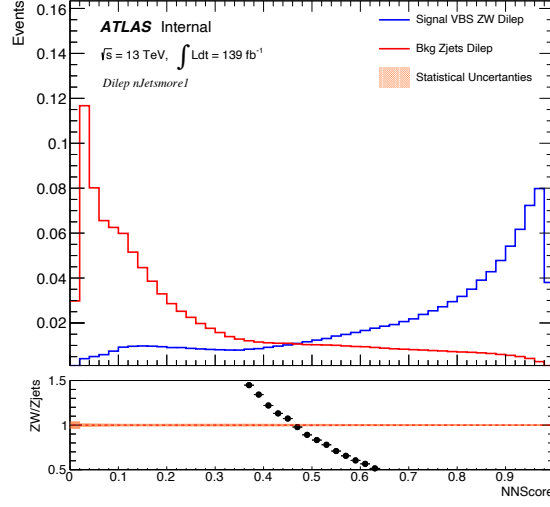


Figure 4.10: Normalized distribution of RNN score for Z +jets background events and ZW signal events. The score peaks around 1 for the latter and around 0 for the former, so that requiring $RNNScore > 0.3$ results in dropping a significant amount of background events while keeping the most part of signal events. The bottom pad displays the ratio between the signal and the background histograms.

where S and B are the total number of signal and background events at a given step of the selection flow, respectively. Equation 4.5 is referred to as total significance.

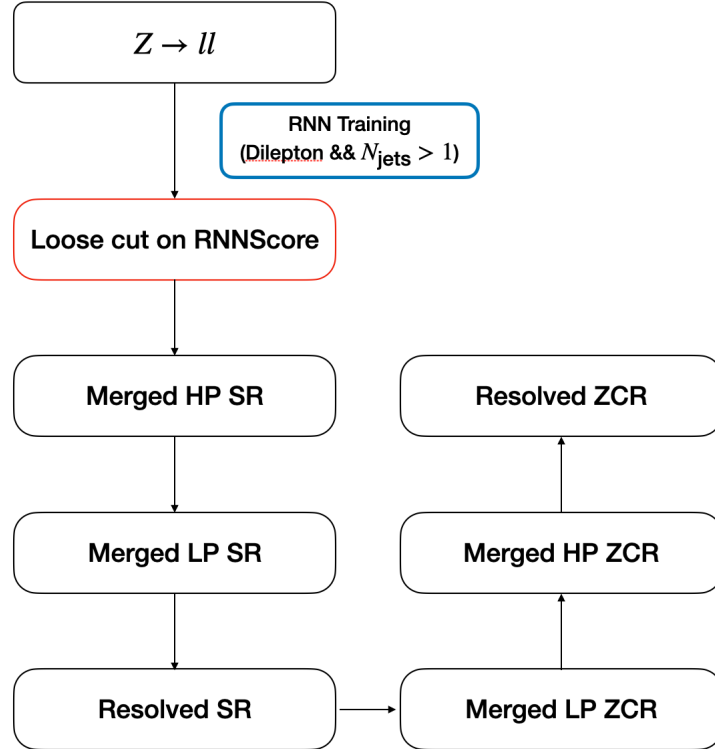


Figure 4.11: Summary of the event categorization in the RNN approach (RNN for VBS categorization).

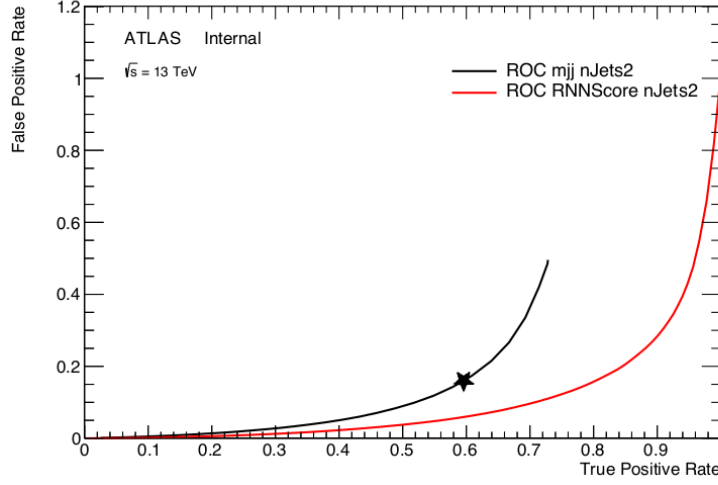


Figure 4.12: RNN ROC for VBS topology selected by the m_{jj}^{tag} cut (black line) and by a RNN score cut (red line). The star point refers to the usual cut $m_{jj}^{\text{tag}} > 400$ GeV.

Figure 4.12 displays the ROC of the RNN approach compared with the ROC of the cut-based approach. In the latter case, the threshold which identifies VBS events is a value of m_{jj}^{tag} . Regardless of the mass cut value, the cut-based TPR is never higher than 0.7, due to VBS signal events with only three small- R jets which are automatically discarded because of the tagging jets search.

The star point on the cut-based ROC refers to the usual cut $m_{jj}^{\text{tag}} > 400$ GeV. As it is made clear from the plot, if we take the FPR value matching the star point, the corresponding TPR value for the RNN curve is higher than the cut-based one.

The following sections illustrate the main contributions with which my work has provided the implementation of the RNN in the current diboson search.

4.4.2.1 Preliminary study of signal leading jet p_T

The signal leading jet p_T cut implemented in the CB approach (namely, $p_T^{j_1} > 40$ GeV) is the result of an optimisation process which is specifically designed for the tagging jets strategy. Switching to the RNN approach makes it necessary to study whether the same cuts fit the new procedure, thus I investigated the efficiency and significance values for different p_T cuts. The signals and total background efficiency trends as the $p_T^{j_1}$ cut varies are displayed in Figure 4.13 for three different values of the RNN score threshold, alongside with both signals total significances. Looking at $p_T^{j_1}$ values around 40 GeV, it is clear that both signal efficiencies and significances are stable, so the usual $p_T^{j_1}$ cut has been kept hereinafter. Obviously, a trade-off between signal efficiency and significance always exists, since tighter cuts applied to the RNN score result in higher significance values but lower signal efficiency values.

4.4.2.2 Signal efficiency and significance with respect to cut-based approach

To assess the impact of the tagging jet search and the m_{jj}^{tag} cut on signal efficiency, I evaluated the latter for different values of the RNN score threshold and compared to the cut-based strategy, alongside with the total and binned significances. As regards the latter, they were calculated on the RNN score distribution in the case of the RNN approach; on the other hand, the m_{jj}^{tag} distributions were exploited in the case of CB approach. This choice appears to be reasonable if one bears in mind that the RNN categorisation replaces the tagging jets strategy. Both merged and resolved signal regions were taken into account.

As it is made clear from Figure 4.14, there is a considerable increase of signal efficiencies with respect to the cut-based values: the gain amounts averaged over different score cut values are displayed in Table 4.9.

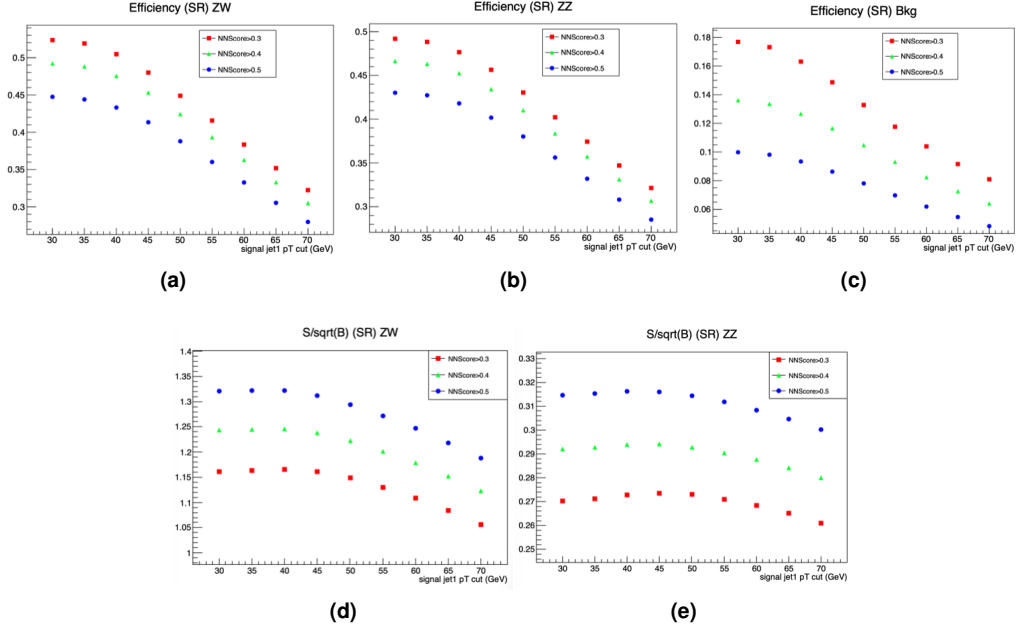


Figure 4.13: Signal and background efficiencies, alongside with total significance for both ZW and ZZ signals, in resolved signal region, varying the leading signal jet p_T cut for different RNN score cuts.

Average signal efficiency increase (%)			
	Merged HP SR	Merged LP SR	Resolved SR
VBS ZZ	100	150	100
VBS ZW	200	300	100

Table 4.9: Average signal efficiency increase (over different RNN score cut values) with respect to the cut-based approach.

Background efficiencies are higher than the cut-based counterparts too, though not high enough to affect the binned significance, which is always higher with respect to the CB approach, as evidenced in Figure 4.15. Here, the blue curve refers to a preliminary RNN training which was performed deploying only ZZ signal events for the "1" category. The binned significances exhibit a quite flat shape as the score cut varies, as it is expected to happen. In fact, the score bins which contribute the most to the binned significance match high values of the score itself (where the maximum lies), so that the asymptotic significance is stable before tight score cuts are applied. This is consistent with the signal efficiency trend, which is rather flat as the score cut varies as well, since the most part of signal events correspond to high score values.

4.4.2.3 The signal jets pairing algorithm in the resolved regime

In the RNN approach, the event selection flow is the same as the cut-based one, with the (fundamental) exception of the search for tagging jets. In particular, the algorithm responsible for the selection of the signal jets pair in the resolved selection is exactly the same, meaning that, among all the available small- R jets, those two such that their invariant mass is the closest one to m_Z or m_W are designated as signal jets candidates. In addition, they must satisfy the p_T and m_{jj} requirements listed in Table 4.6 for the resolved SR to be defined as signal jets. Taking its peculiar nature into account, such a pairing algorithm will be hereinafter referred to as "close V", for the sake of simplicity.

From a physical point of view, the close V algorithm is the most intuitive one it can be thought of when choosing a strategy to identify particles resulting from another (known) particle decay. However, this very algorithm induces a bias in the distribution of the signal jets invariant mass m_{jj} , especially when evaluated for background events. More specifically, even though a given jets pair is not related to the decay of a boson,

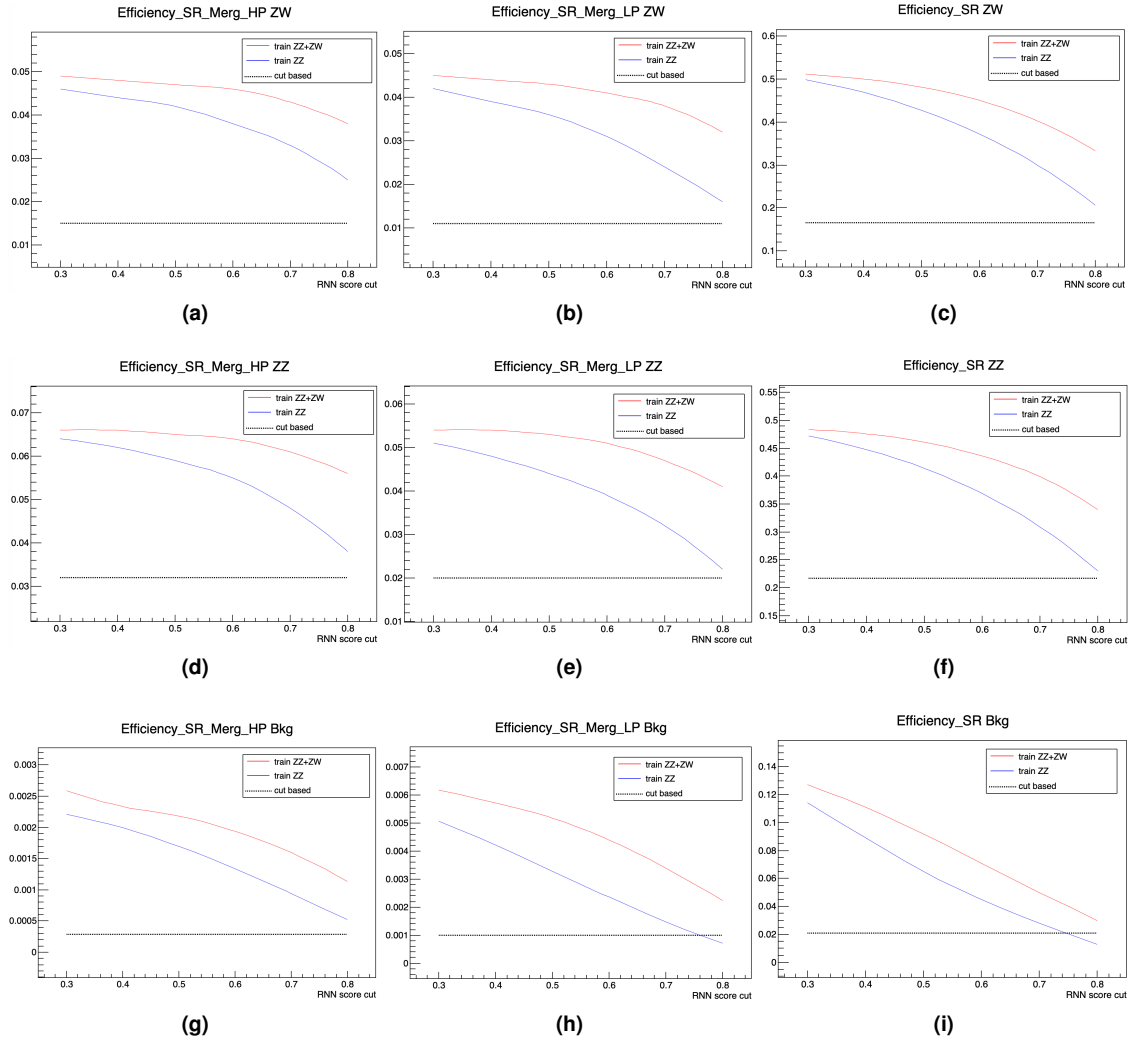


Figure 4.14: Efficiencies of ZW (first row) and ZZ (second row) signals and the total background (third row) in merged (first two columns) and resolved (third column) signal regions, varying the RNN score cut (red line) versus the cut-based counterparts (grey dotted line). The blue lines refer to a previous training of the RNN performed using ZZ signal only.

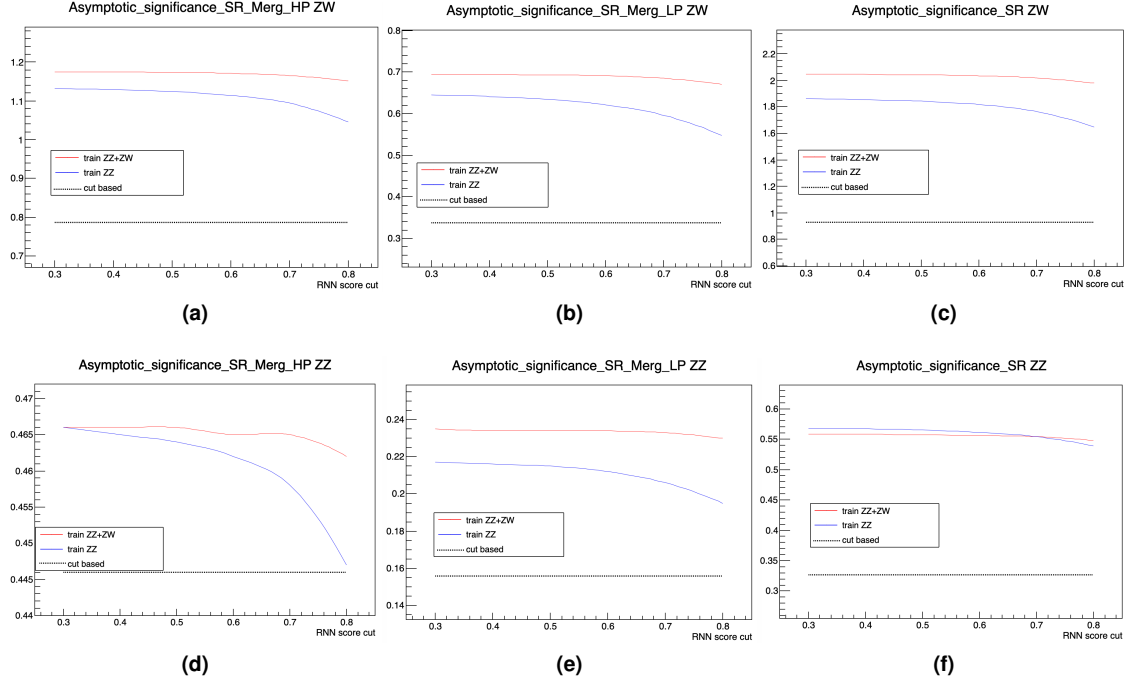


Figure 4.15: Asymptotic significances of ZW (first row) and ZZ (second row) signals evaluated on the RNN score distributions in merged (first two columns) and resolved (third column) signal regions, varying the RNN score cut (red line) versus the cut-based counterparts, evaluated on m_{jj}^{tag} distributions (grey dotted line). The blue lines refer to a previous training of the RNN performed using ZZ signal only.

it can accidentally have an invariant mass which is dangerously close to m_Z or m_W . In the case of signal events (meaning that VBS with a hadronically decaying boson actually occurs), a "fake" signal pair can be chosen over the "real" one, perhaps because one jet of the "real" pair is so poorly reconstructed that the pair turns out to have an invariant mass which differs from m_Z or m_W by a significant amount; consequently, the expected peak in the m_{jj} distribution around m_Z or m_W is artificially enhanced (and so is the signal efficiency). In the case of background events, m_{jj} distribution should not exhibit any peaks at all, yet the close V algorithm induces artificial maximums around $m_{Z/W}$, as illustrated in Figure 4.16. This of course negatively affects significance.

As it is shown in Figure 4.17, truth Monte Carlo simulations reveal that signal jets are mostly central, i.e. they satisfy the condition $|\eta_j| < 2.5$. My first attempt to reduce the number of "fake" signal jets pairs which end up in the resolved signal region consisted in excluding non-central (or forward) signal jets. This request resulted in dropping 10% of data events in the resolved preselection (namely, just before the mass window requirement which defines the resolved signal region).

Intuitively, one would expect a more flat distribution of m_{jj} for both background and signal events, since the "fake" signal jets pairs which enhance the artificial peaks should be removed. Indeed, this is what can be seen in Figure 4.18, where the distribution of m_{jj} for signal and background events is displayed in resolved preselection (without any cut applied to the RNN score).

Looking at the efficiencies in Table 4.10, one can see that the percentage decrease of the background efficiency coming from dropping the forward jets from the signal jets list is consistent with the 10% loss of data yields mentioned above. The decrease in the asymptotic significance (calculated on m_{jj}) is quite stable moving from the resolved preselection to the resolved signal region (i.e. it is not affected by the mass window requirement on m_{jj}). However, there exists a significant efficiency loss (more than 10%) when it comes to signals in the resolved signal region with respect to the decrease that is observed in the resolved preselection. This could be due to several factors:

- in some cases, one (or both) signal jet is actually forward;
- both signal jets are central, but one of them is so poorly reconstructed that makes the pair fail to meet

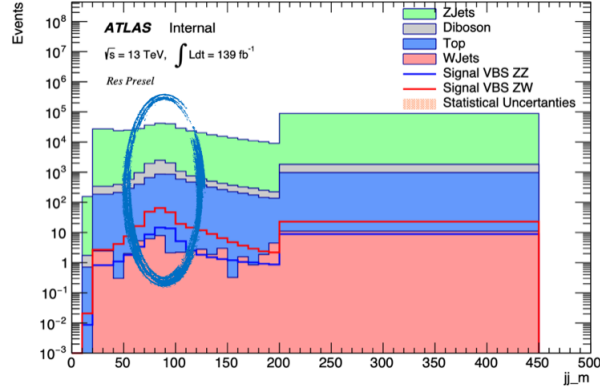


Figure 4.16: m_{jj} distributions for signal and background events in resolved preselection (i.e. just before the resolved signal region mass window requirement). The close V algorithm induces a bias in both signal and background distributions by giving rise to artificial maximums around $m_{Z/W}$ as it has been highlighted in the plot.

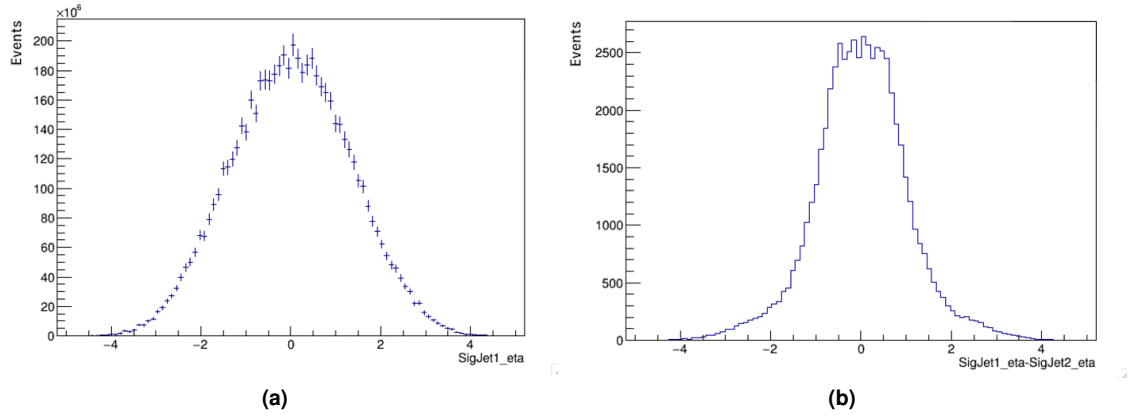


Figure 4.17: (a): η distribution for the truth signal leading jet is mainly contained in the $|\eta| < 2.5$ interval, meaning that the leading signal jets are mostly central. (b): $\eta_1 - \eta_2$ distribution for truth signal jets is centred around zero, meaning that the subleading signal jets are mostly central as well.

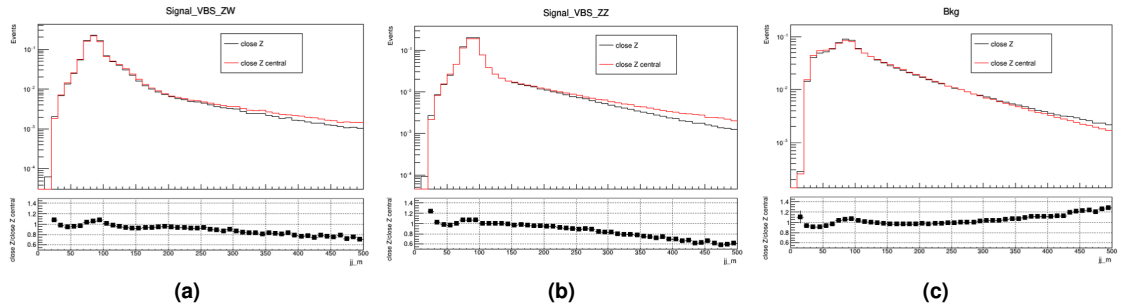


Figure 4.18: m_{jj} distribution for signal and background events before (black line) and after (red line) removing forward jets from the signal jets, in resolved preselection. The bottom pads display the ratios between the histograms before and after removing the forward jets.

the mass requirement of the signal region;

- since the resolved selection is tested after the merged selection, one of the jets picked as signal could actually be a large- R jet candidate that did not pass the merged selection cuts.

In order to investigate the last two bullets, the signal region mass window was widened ($60 \text{ GeV} < m_{jj} < 110 \text{ GeV}$) and the absence of large- R jets was required. However, the efficiency loss does not seem to be significantly affected.

Resolved Preselection			
	Close V	Close V central	Decrease (%)
ϵ_B	0.65	0.58	10
ϵ_{ZW}	0.86	0.82	5
ϵ_{ZZ}	0.85	0.80	5
$\sigma_{\text{asy},ZW}(m_{jj})$	1.05	1.02	2
$\sigma_{\text{asy},ZZ}(m_{jj})$	0.24	0.23	4
Resolved SR			
	Close V	Close V central	Decrease (%)
ϵ_B	0.20	0.18	12
ϵ_{ZW}	0.53	0.48	10
ϵ_{ZZ}	0.50	0.44	12
$\sigma_{\text{asy},ZW}(m_{jj})$	0.96	0.93	3
$\sigma_{\text{asy},ZZ}(m_{jj})$	0.22	0.21	6
Resolved SR (60-110), no large- R jets			
	Close V	Close V central	Decrease (%)
ϵ_B	0.20	0.18	12
ϵ_{ZW}	0.42	0.38	10
ϵ_{ZZ}	0.35	0.30	13
$\sigma_{\text{asy},ZW}(m_{jj})$	0.77	0.74	4
$\sigma_{\text{asy},ZZ}(m_{jj})$	0.16	0.15	7

Table 4.10: Signals and background efficiencies, together with the asymptotic significances (computed on m_{jj} distribution) before (close V) and after (close V central) removing the forward jets from the signal jets list, in the resolved preselection, resolved signal region and in a "widened" resolved signal region with no large- R jets in the event. The fourth column shows the yields downsize.

Applying cuts to the RNN score does not seem to reduce the efficiency loss either, as it is made clear in Figure 4.19: the gap between the signal efficiencies before and after removing the forward jets is still sizeable and does not seem to depend on the RNN score threshold.

In conclusion, the close V algorithm is still under investigation. An alternative signal jets pairing algorithm will be discussed in section 4.4.3.2.

4.4.2.4 RNN score tight cuts

As referred to in the previous paragraphs, the RNN has been initially implemented in order to replace the tagging jets search as a means of recognizing VBS topology: the neural network training was performed at dilepton level (when at least two small- R jets were found) and a loose cut was intended to be applied to the RNN score. Nevertheless, the latter is only one approach out of many which could virtually lead to a sizeable signal efficiency enhancement. Bearing in mind the issues arising from the close V algorithm implementation that have been discussed in section 4.4.2.3, one could think to make a different use of the RNN: instead of defining the resolved signal region by mass cuts, a high value of the RNN score can be required for the events which pass the resolved pre-selection, the latter being defined just before the resolved signal region mass window cut is applied. In this way, both the tagging jets search and the mass window requirement functions are taken care of by the RNN.

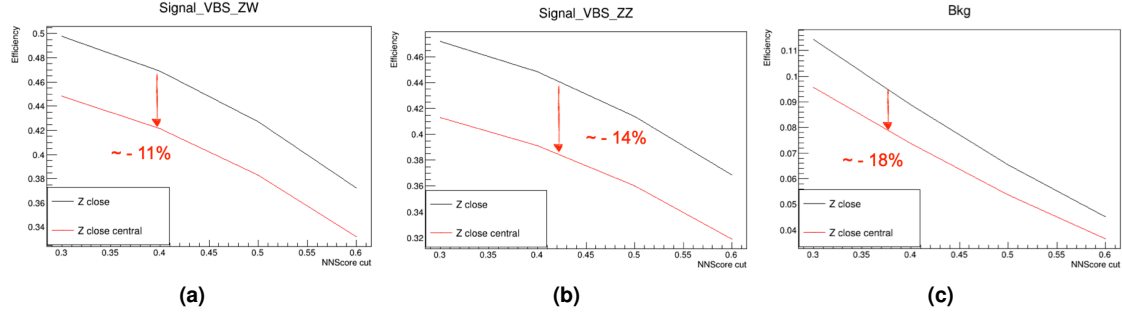


Figure 4.19: Signal and background efficiencies in resolved signal region before (black line) and after (red line) removing forward jets from the signal jets list, as the RNN score threshold varies.

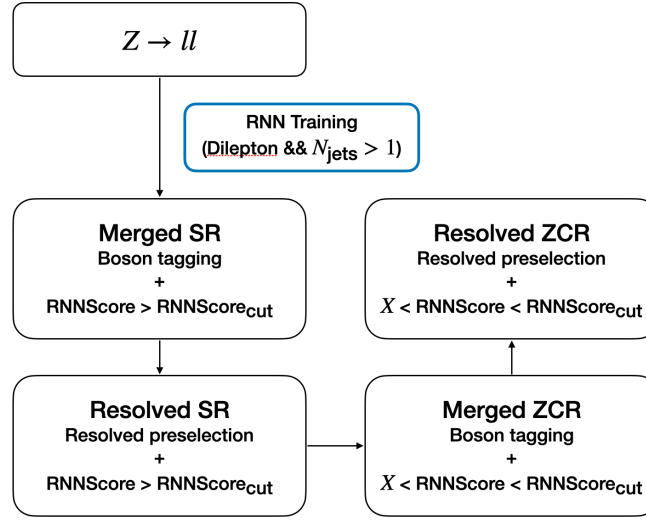


Figure 4.20: Summary of the event categorization in the RNN approach, without the search for tagging jets nor the mass windows requirement of the cut-based resolved and merged signal regions.

The new selection flow is illustrated in Figure 4.20. For the sake of simplicity, only one merged signal region is defined.

I took three different score tight cuts into consideration:

- $RNNscore_{cut}=0.7$;
- $RNNscore_{cut}=0.8$;
- $RNNscore_{cut}=0.9$.

In fact, in order to obtain a background efficiency of the same magnitude as in the cut-based signal regions, a RNN score value higher than 0.9 was needed for both the resolved and the merged (inclusive) selections (Figure 4.21). The distributions of signal jets m_{jj} , the small- R jets multiplicity N_{jets} and the diboson system invariant mass m_{VV} before and after applying the above mentioned score cuts are displayed in Figure 4.22 for the resolved case.

As one can infer by looking at Table 4.11 and Figure 4.21, applying tight cuts to the RNN score results in nearly tripled signal efficiencies with respect to the cut-based values, for both the resolved and the merged signal regions. The asymptotic significances, which are now computed on the diboson system invariant mass m_{VV} (in order to have a fair comparison with the CB approach), are enhanced as well.

With plans to define new resolved and merged signal regions by applying tight cuts on the RNN score, I performed a preliminary study to identify new suitable definitions of the ZCRs as well, taking the resolved case into account.

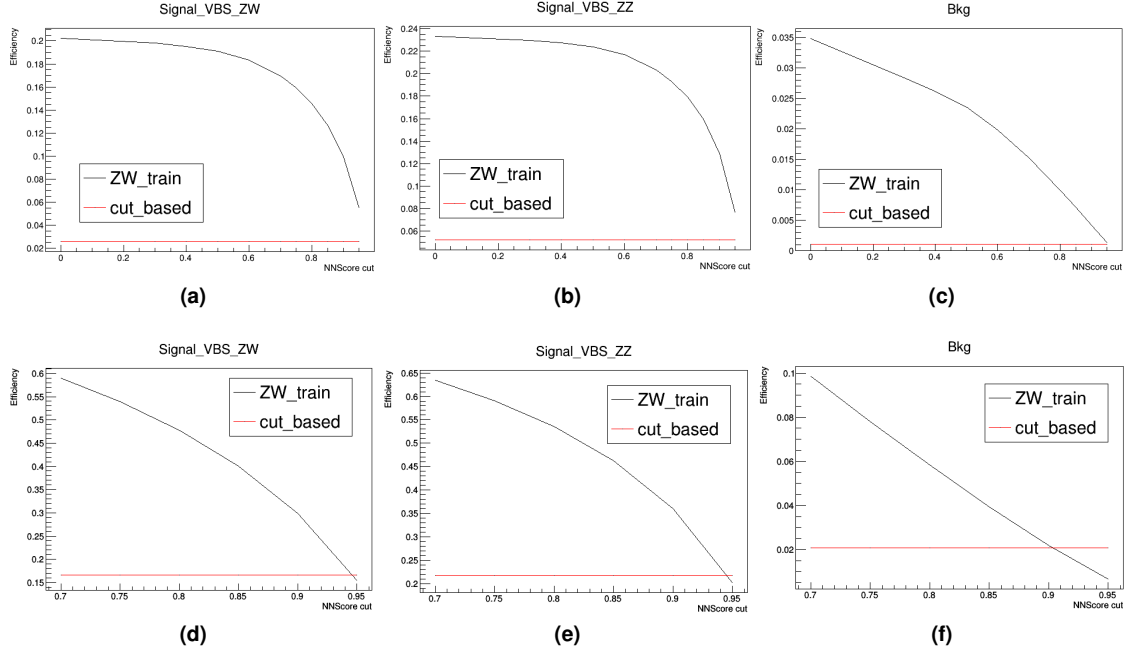


Figure 4.21: Signal and background efficiency in merged (a) - (c) and resolved (d) - (f) signal regions for CB approach (red line) and the RNN approach (black line), as the RNN score threshold varies.

	Resolved SR				Merged (inclusive) SR			
	CB approach	RNNscore>0.7	RNNscore>0.8	RNNscore>0.9	CB approach	RNNscore>0.7	RNNscore>0.8	RNNscore>0.9
ϵ_{ZW}	0.17	0.59	0.48	0.30	0.03	0.17	0.14	0.10
ϵ_{ZZ}	0.22	0.63	0.54	0.36	0.05	0.20	0.18	0.13
$\sigma_{asy,ZW}(m_{VV})$	0.93	1.48	1.57	1.61	0.64	1.06	1.13	1.19
$\sigma_{asy,ZZ}(m_{VV})$	0.30	0.40	0.44	0.48	0.33	0.31	0.34	0.38

Table 4.11: Signal efficiencies and asymptotic significances (evaluated on the diboson system invariant mass m_{VV}) in resolved and merged (inclusive) signal regions, both for the CB approach and as the high RNN score threshold varies.

In general, a proper control region is characterized by a trade-off between:

- experimental data being (for the most part) distributed in the control region accordingly to the background (i.e., a small quantity of signal events)
- the control region being similar to the signal region in kinematics terms

so that it is possible to extract the number of background events in the signal region on the basis of what is observed in the control region.

In the cut-based approach, the resolved ZCR is defined by the same selections of the resolved SR, with the exception of the mass window cut, which is reversed:

$$54 \text{ GeV} < m_{jj} < 64 \text{ GeV} \quad \text{or} \quad m_{jj} > 106 \text{ GeV}.$$

As for the RNN approach we are discussing, the resolved signal region is defined by the condition $\text{RNNscore} > \text{RNNscore}_{\text{cut}}$, so that reversing this selection results in requiring that

$$X < \text{RNNscore} < \text{RNNscore}_{\text{cut}}$$

where X is a score value to be properly chosen taking two main factors into consideration:

- X should be low enough to let a sizeable amount of background events in the control region, in order to reduce statistical uncertainty;
- X should be high enough to let the control region "resemble" the signal region as much as possible.

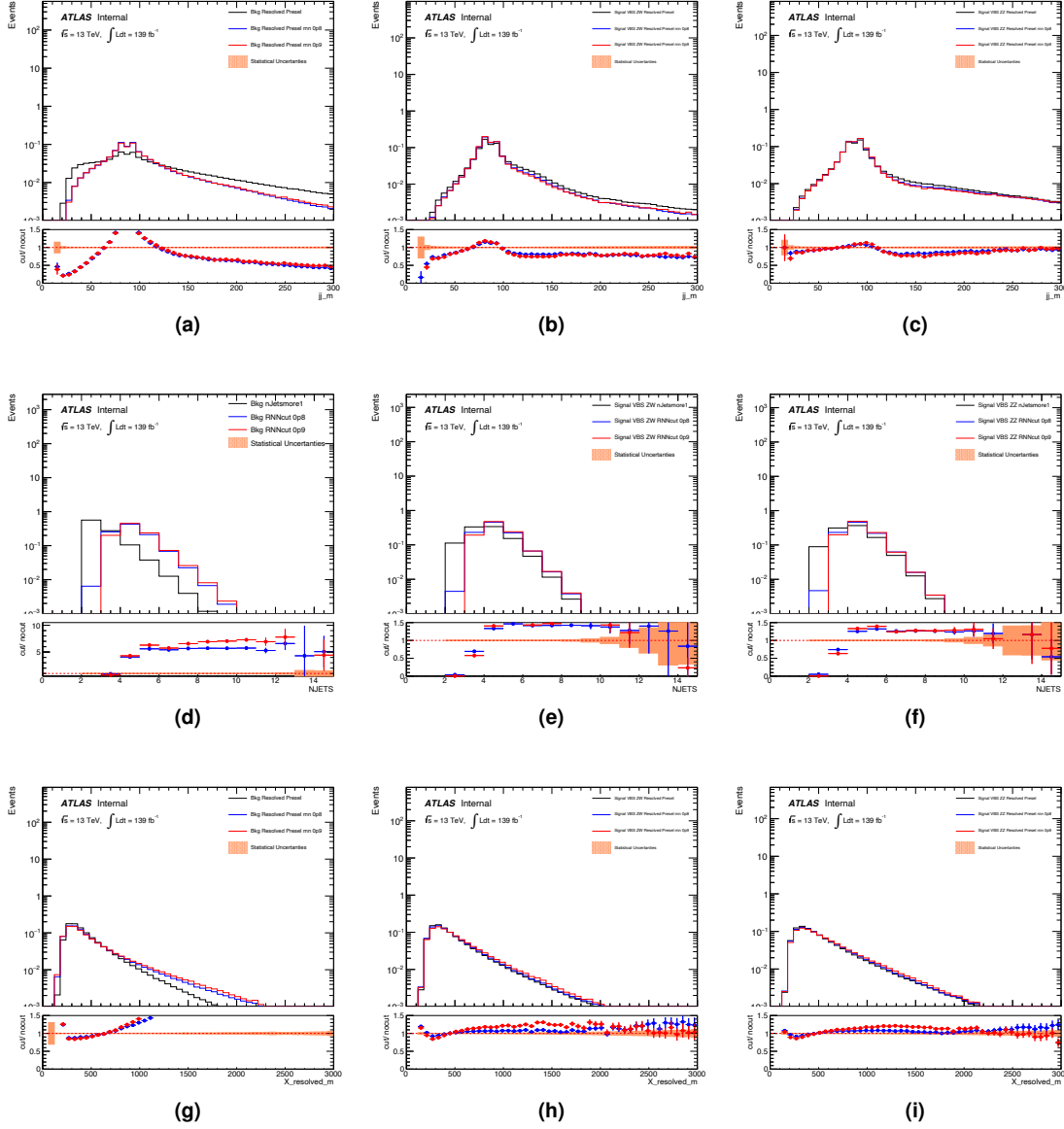


Figure 4.22: Distribution before (black) and after applying cuts on the RNN score (0.8: blue, 0.9: red) of m_{jj} (a) - (c), the number of small-R jets (d) - (f) and m_{VV} (g) - (i) in resolved preselection, for background (first column), ZW signal (second column) and ZZ signal (third column) events. The bottom pads display the ratios between the histograms after and before applying the cuts to the score.

Again, the choice of X turns out to be a trade-off. Depending on the value of $\text{RNNScore}_{\text{cut}}$, different values of X were tested (Table 4.12). The "purity" of the ZCR refers to the fraction of Z +jets yields with respect to the total background yields in the ZCR. In all three cases $\text{RNNScore}_{\text{cut}} = 0.9, 0.8, 0.7$, the purity diminishes as X increases, since the score window defining the ZCR itself shrinks. Nevertheless, the purity for all X values remains close to the cut-based purity. The ratio between the Z +jets yields in the ZCR and the Z +jets yields in the resolved SR was estimated as well: a suitable control region would hopefully contain at least the same number of background events as the one expected in the signal region, so the aforementioned ratio should be at least 1. Finally, the total significances were evaluated: the ideal X should be such that σ is as low as possible.

		Z +jets yields	ZCR purity	Z +jets yields (ZCR/SR)	σ_{ZZ}	σ_{ZW}
Cut-based		92179	0.96	1.49	0.16	0.53
$\text{RNNScore}_{\text{cut}} = 0.9$	$X = 0.4$	656966	0.96	10.14	0.17	0.79
	$X = 0.5$	507895	0.96	7.84	0.18	0.82
	$X = 0.6$	363475	0.95	5.61	0.19	0.84
	$X = 0.7$	230691	0.95	3.56	0.19	0.83
	$X = 0.8$	109594	0.95	1.69	0.18	0.74
		Z +jets yields	ZCR purity	Z +jets yields (ZCR/SR)	σ_{ZZ}	σ_{ZW}
Cut-based		92179	0.96	1.49	0.16	0.53
$\text{RNNScore}_{\text{cut}} = 0.8$	$X = 0.4$	483536	0.96	2.77	0.10	0.53
	$X = 0.5$	366339	0.96	2.10	0.11	0.53
	$X = 0.6$	239078	0.95	1.37	0.11	0.52
	$X = 0.7$	116125	0.95	0.67	0.10	0.44
		Z +jets yields	ZCR purity	Z +jets yields (ZCR/SR)	σ_{ZZ}	σ_{ZW}
Cut-based		92179	0.96	1.49	0.16	0.53
$\text{RNNScore}_{\text{cut}} = 0.7$	$X = 0.4$	367411	0.96	1.24	0.07	0.35
	$X = 0.5$	250214	0.96	0.85	0.07	0.34
	$X = 0.6$	122953	0.96	0.42	0.06	0.28

Table 4.12: Resolved ZCR summary for the cut-based approach and the RNN approach, in correspondence of $\text{RNNScore}_{\text{cut}}=0.9, 0.8, 0.7$ and different values of the lower score value X . The fourth column displays the ratio between the Z +jets yields in the ZCR and the Z +jets yields in the SR.

Figure 4.23 shows the distribution of m_{jj} , m_{jj}^{tag} , the number of small- R jets (N_{jets}), m_{VV} , η and p_T of the first small- R jets in both the cut-based resolved SR and resolved ZCR, in reference to Z +jets events. The same variables are displayed in Figure 4.24 for the RNN approach in the case of $\text{RNNScore}_{\text{cut}} = 0.8$ and $X = 0.7, 0.6, 0.5, 0.4$. The tagging jets are defined in the same way as the cut-based approach, but they are not used in any selection. The ratio plots refer to the ZCR/SR ratio, so for the ZCR to be acceptable, those very plots should be as flat as possible, since they are concerned with Z +jets events.

Taking the values displayed in both Tables 4.11 and 4.12 into account, the 0.9 score cut seems to heavily affect the signal efficiency (with respect to the 0.8 and 0.7 cuts) without really bringing benefits to the ZCR definitions. On the other hand, a suitable choice for $\text{RNNScore}_{\text{cut}} = 0.8$ could be $X = 0.5$ or $X = 0.4$. In fact, the ZCR purity is always about 95% ÷ 96%, so that the trade-off needs to be found only between the expected significance and the fraction of Z +jets yields in the ZCR with respect to the Z +jets yields in the signal region. Similarly, $X = 0.4$ could be a reasonable option as for $\text{RNNScore}_{\text{cut}} = 0.7$.

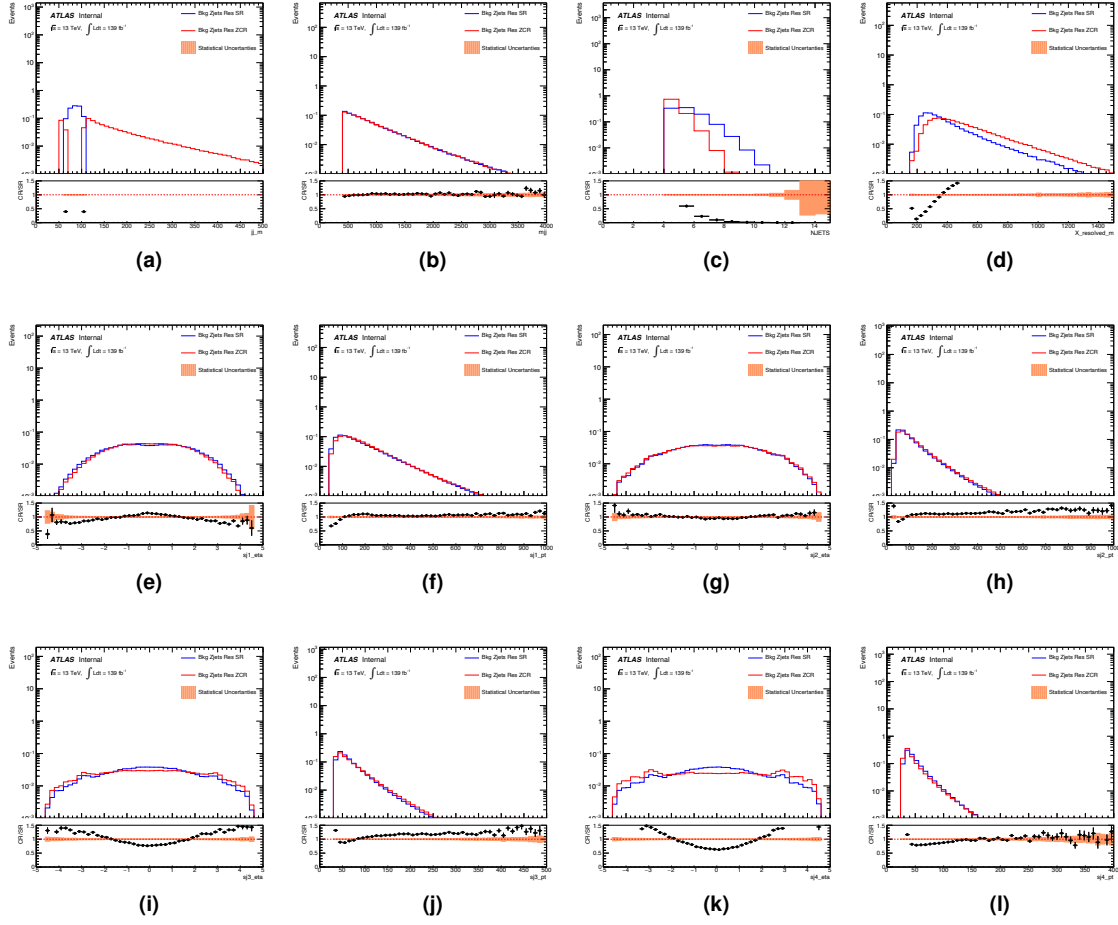


Figure 4.23: Distribution in cut-based resolved SR (blue line) and resolved ZCR (red line), for Z +jets events, of m_{jj} (a), m_{jj}^{tag} (b), the number of small- R jets (c), m_{VV} (d), η and p_T of the first four small- R jets (e) - (l). The bottom pads display the ratio between the histograms in the control region and in the signal region.

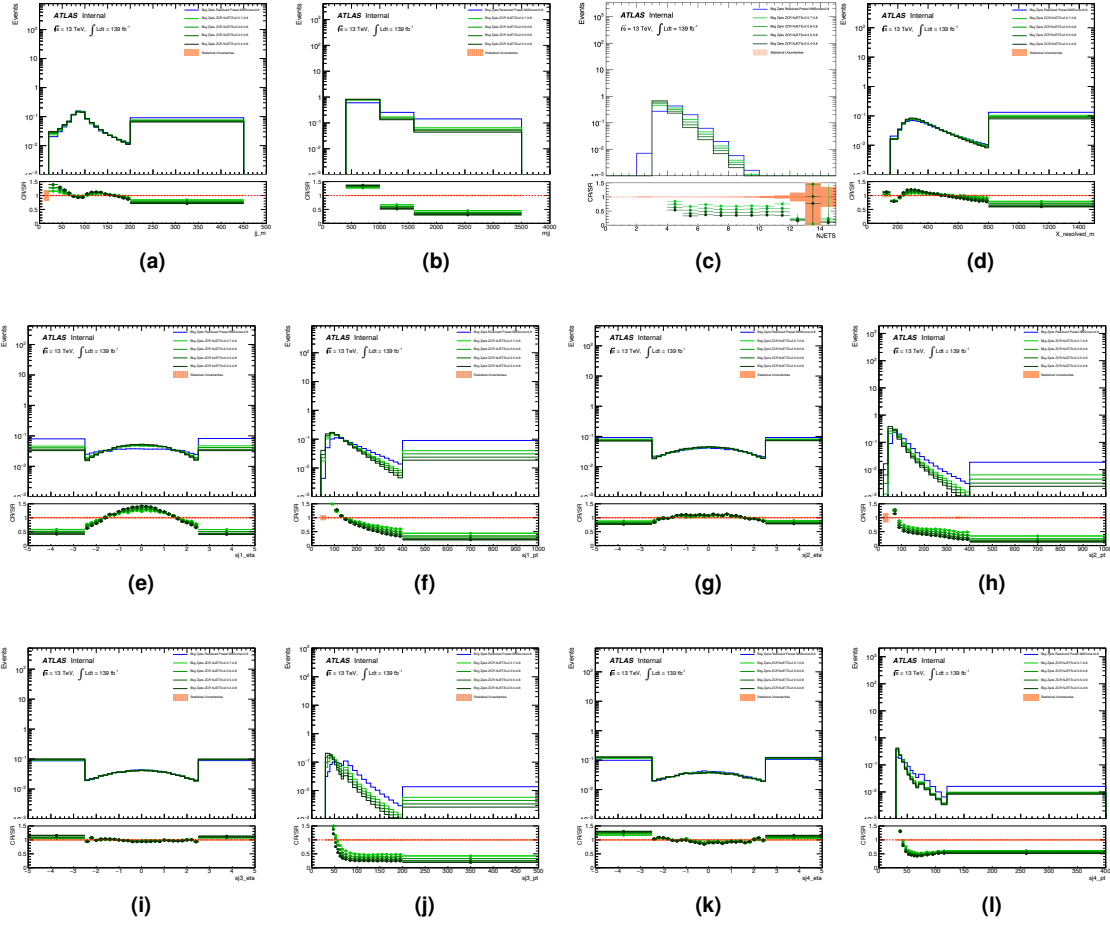


Figure 4.24: Distribution in resolved preselection after 0.8 score cut (blue line) and resolved ZCR varying X (green lines), for Z +jets events, of m_{jj} (a), m_{jj}^{tag} (b), the number of small- R jets (c), m_{VV} (d), η and p_T of the first four small- R jets (e) - (l). The bottom pads display the ratio between the histograms in the control region and in the signal region.

4.4.3 Training in signal regions

As already pointed out in the previous sections, when trained at dilepton level (if at least two small- R jets are found in the events), the RNN provides the electroweak diboson production search with higher signal efficiency and higher significance in the signal regions with respect to the cut-based approach. In other words, the RNN achieves satisfactory results by fulfilling a classification task, namely identifying VBS topology.

Nonetheless, it will be shown in this section that even better performances can be obtained if the RNN is trained in the resolved and merged signal regions, with the aim to maximise the signal-background separation. Instead of applying cuts to the RNN score in order to classify the events, the score itself can be used as final discriminant, i.e. as a variable to whose distribution a fit can be performed (in the signal region) in order to extract the signal strength μ (see Chapter 5).

For the purpose of assessing the signal efficiency and the expected significance enhancement, a comparison with the previous run cut-based approach involving a Boosted Decision Tree (BDT) score as final discriminant [31] is provided (the cut-based asymptotic significance is therefore calculated on the BDT score). In particular, only the 2015-2016 Monte Carlo samples (i.e. a 36 fb^{-1} integrated luminosity) are considered in the case of the RNN approach, to have a fair comparison with the previous analysis. The expected binned significances are evaluated on the RNN score.

Since the resolved and merged signal regions differ by the type of jets collection from which the signal jets are selected (small- R jets or large- R jets, respectively), the RNN new training performed in such regions results in providing two different scores: a resolved-case score and a merged-case score.

As regards the resolved-case RNN score, it is the outcome of a training performed using the usual four-momentum variables plus the number of tracks of the small- R jets as input features. The RNN deals with variable-length inputs, since the first small- R jets are considered for each training event, up to a maximum of six jets. The events involved with the training are those passing the cut-based resolved signal region selection (Table 4.6) with the only exception of the search for tagging jets, i.e. those events satisfying the conditions:

- 2 leptons passing the dilepton selection;
- no large- R jets or large- R jets which fail both the merged HP and LP signal regions requirements;
- at least 2 small- R jets;
- signal jets:
 - ◊ $\min |m_{jj} - m_{Z/W}|$;
 - ◊ $64 \text{ GeV} < m_{jj} < 106 \text{ GeV}$;
 - ◊ $p_T^{j1} > 40 \text{ GeV}, p_T^{j2} > 20 \text{ GeV}$.

The training is performed in a supervised mode: each training event is known to be signal-like (whether ZZ or ZW signal) or background-like (whether Z +jets, diboson or top background), so that the network is able to discriminate between these two categories.

The RNN score for ZZ and ZW signals, Z +jets, diboson and top background processes is illustrated in Figure 4.25. With respect to the training after dilepton selection, the background shapes maximums seem to be closer to zero, meaning that the resolving power of the network is a little enhanced, while the performance of the neural network in terms of the area under the ROC curve remains solid, as well as the expected binned significances (Figure 4.26).

As for the training in the merged signal region, it is performed using events satisfying the conditions:

- 2 leptons passing the dilepton selection;
- at least one large- R jet;
- signal jet:
 - ◊ p_T -leading;
 - ◊ 50 % (HP) or 80 % (LP) WP boson tagging $D_2^{\beta=1}$, n_{track} and $m_{W/Z}$;

- ◊ $p_T^J > 200 \text{ GeV}, |\eta_J| < 2;$
- ◊ $\Delta R(l_{1,2}, J) > 1.0;$

(namely, the events passing the HP or the LP merged cut-based signal region selections, with the only exception of the search for tagging jets). Again, each training event is labelled as signal- or background-like. Taking the same variables employed in the resolved case training as input features, a sizeable improvement in the network ability to recognize background events as such is achieved, as it can be seen in Figure 4.27. In consequence, the performance in terms of both the area under the ROC curve and the expected binned significances is increased with respect to the dilepton training (Figure 4.28).

In order to further improve the merged performance by taking the information about the signal large- R jet into account, six additional variables were fed to the RNN as input features, alongside with the four-momentum variables and the number of track of the first six small- R jets. These additional features are:

- the four-momentum variables of the large- R jet;
- the number of tracks of the large- R jet;
- the $D_2^{\beta=1}$ of the large- R jet.

Consequently, the RNN trained in the inclusive (HP + LP) merged signal region finally allows for a variable-length input (events with at least one large- R jet and up to six small- R jets) and a total of eleven input features for each event.

The distributions of the additional input features regarding the large- R jet for signal and background events are illustrated in Figure 4.29, while the score shapes for signal and background events, as well as the ROC curves and the expected binned significances with respect to the "standard" training in the merged signal region (i.e, without employing the large- R jet variables as input features) are shown in Figures 4.30 and 4.31.

Not only does this new training set-up improve the inclusive (HP + LP) merged analysis, it also increases the network performance in the merged high purity and low purity signal regions separately (Figures 4.31c - 4.31f). Table 4.13 summarizes the improvement obtained with respect to the dilepton training.

36 fb^{-1}	RNN (training at dilepton && $N_{\text{jets}} > 1$)				RNN (training in SRs, improved merged)				% increase
	Merged HP SR	Merged LP SR	Resolved SR	Combined	Merged HP SR	Merged LP SR	Resolved SR	Combined	Combined
$\sigma_{\text{asy}, ZZ/ZW}$	0.93	0.50	1.69	1.99	1.45	0.61	1.72	2.34	18

Table 4.13: Expected binned significances (considering both ZZ and ZW signals together) comparison between the RNN training performed at dilepton selection and the RNN training performed in the merged and resolved signal regions. The significance values are computed on the RNN score. The value combined over the three signal regions is estimated by means of the square root of the squaring sum of the three separate significances.

The most interesting result comes from the comparison between the cut-based approach and the RNN new configuration. The cut-based binned significance values presented in Table 4.14 are calculated on the score of the BDT.

36 fb^{-1}	Cut-based				RNN (training in SRs, improved merged)				% increase
	Merged HP SR	Merged LP SR	Resolved SR	Combined	Merged HP SR	Merged LP SR	Resolved SR	Combined	Combined
Yields ZZ/ZW	9.7	6.2	111.1	127.0	28.0	23.0	405.5	456.5	270
$\sigma_{\text{asy}, ZZ/ZW}$	1.06	0.35	0.67	1.34	1.45	0.61	1.72	2.34	75

Table 4.14: Signal yields and binned significances (considering both ZZ and ZW signals together) comparison between the cut-based approach and the RNN approach. The cut-based significance is evaluated on the score of the BDT, while the RNN counterpart is computed on the RNN score. As for the RNN approach, the significance combined over the three signal regions is estimated by means of the square root of the squaring sum of the three separate significances.

As for the full 139 fb^{-1} integrated luminosity, Figure 4.32 shows the score distributions in the merged HP,

merged LP and resolved signal regions. The bottom pads illustrate the signal-to-background ratio, as well as the expected binned significance (for ZZ and ZW together) calculated on the score. The following values are found:

139 fb^{-1}	Merged HP SR	Merged LP SR	Resolved SR	Combined
$\sigma_{\text{asy}, ZZ/ZW} (\text{RNN Score})$	2.58	1.20	3.36	4.40

Table 4.15: Merged HP, merged LP and resolved signal region expected values of the asymptotic significance (for the ZZ and ZW signals together) evaluated on the RNN score, when the RNN is trained in the signal regions.

The signal efficiency benefits from the RNN training in the signal regions as well as the expected significance.

In the cut-based resolved selection, at least four small- R jets are required for the events to pass the resolved signal region. As already explained in section 4.4, the RNN approach, while avoiding the search for tagging jets, makes it possible to retrieve signal events with only three small- R jets in the resolved selection. Table 4.16 displays the amount of such events which are recovered with the RNN trained in the signal regions. About 30(40)% of $ZZ(ZW)$ signal events with $N_{\text{jets}} = 3$ is recovered, positively affecting the signal efficiency.

139 fb^{-1}	Dilepton	Dilepton (3 jets)	Dilepton (3 jets) $\div$ Dilepton	Resolved SR	Resolved SR, 3 jets	Resolved SR (3 jets) $\div$ Resolved SR	Efficiency (3 jets)
VBS ZZ	524.8	167.2	0.32	265.24	57.03	0.22	0.34
VBS ZW	2435.8	833.0	0.34	1230.0	350.9	0.27	0.42

Table 4.16: Yields and efficiencies of ZZ and ZW signals in the RNN approach, with the network trained in the merged and resolved signal regions. At dilepton level, about 30% of signal events have $N_{\text{jets}} = 3$ (for ZZ and ZW signals separately), meaning that they are discarded in the cut-based approach. As for the resolved signal region, about 25% of signal events have $N_{\text{jets}} = 3$ (for ZZ and ZW signals separately), so that the percentage of three jets signal events recovered in the signal region with respect to the three jets signal events at dilepton level is about $30 \div 40\%$.

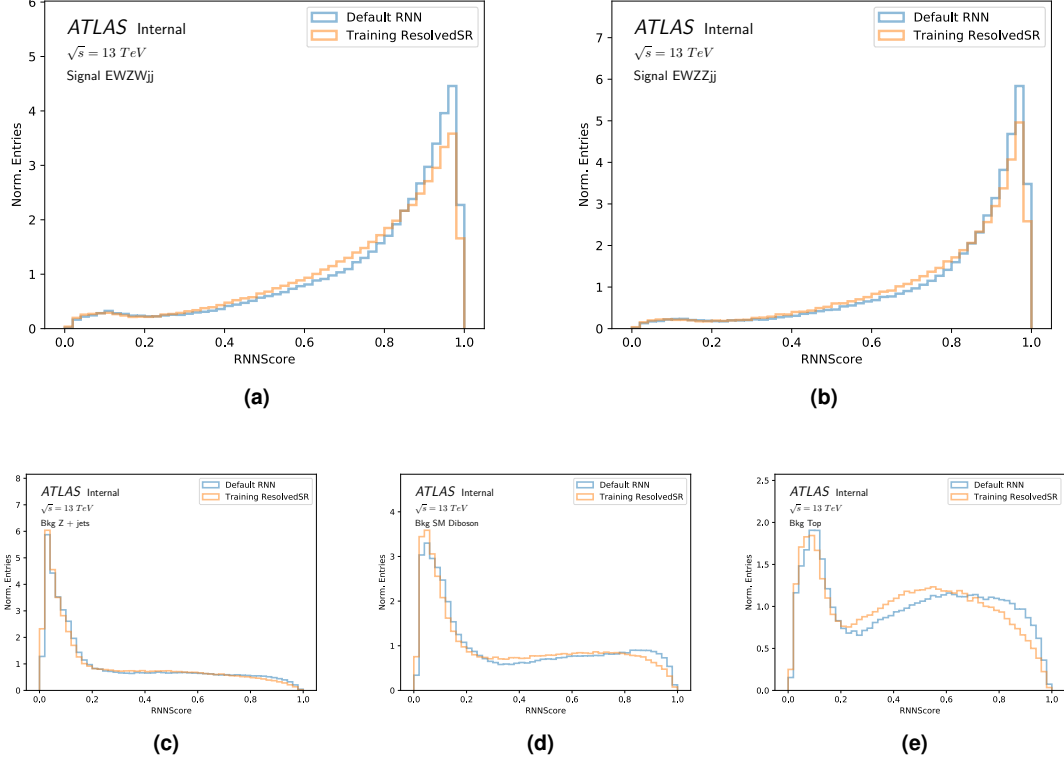


Figure 4.25: Distribution of the RNN score after the dilepton training (blue line) and the training in the resolved signal region without the tagging jets selections (orange line), for signals and Z +jets, diboson and top backgrounds.

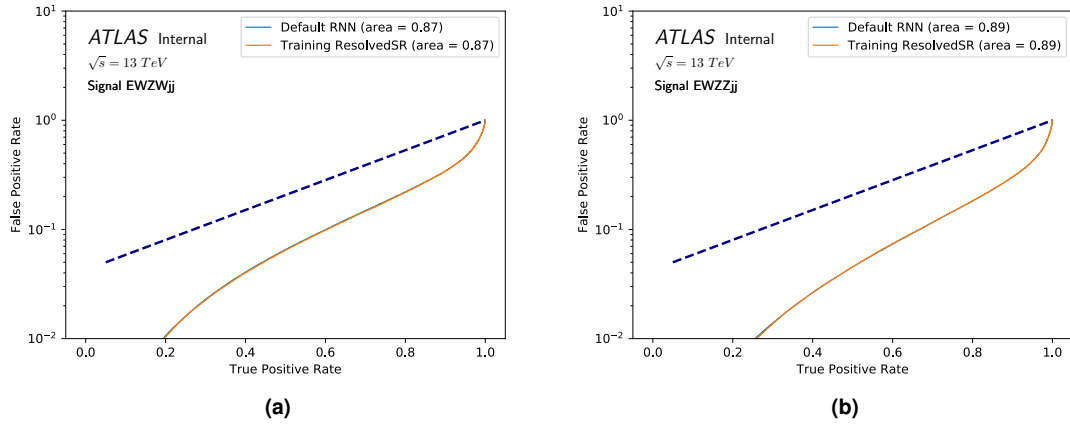


Figure 4.26: ROC curves after the dilepton training (azure line) and the training in the resolved signal region without the tagging jets selections (orange line), for ZW and ZZ signals.

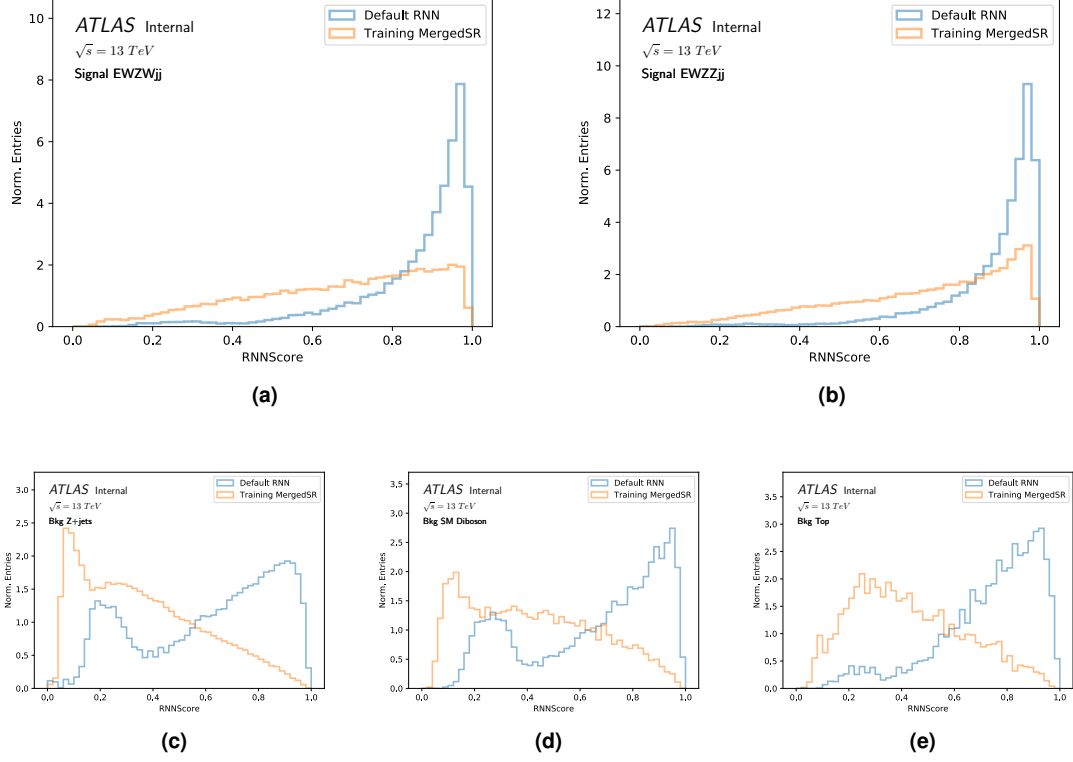


Figure 4.27: Distribution of the RNN score after the dilepton training (blue line) and the training in the merged inclusive (HP + LP) signal region without the tagging jets selections (orange line), for signals and Z+jets, diboson and top backgrounds.

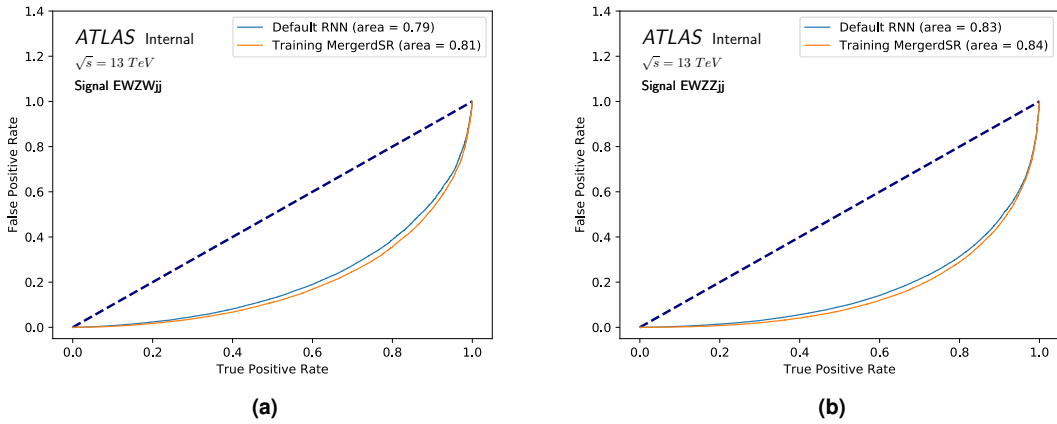


Figure 4.28: ROC curves after the dilepton training (azure line) and the training in the merged inclusive (HP + LP) signal region without the tagging jets selections (orange line), for ZW and ZZ signals.

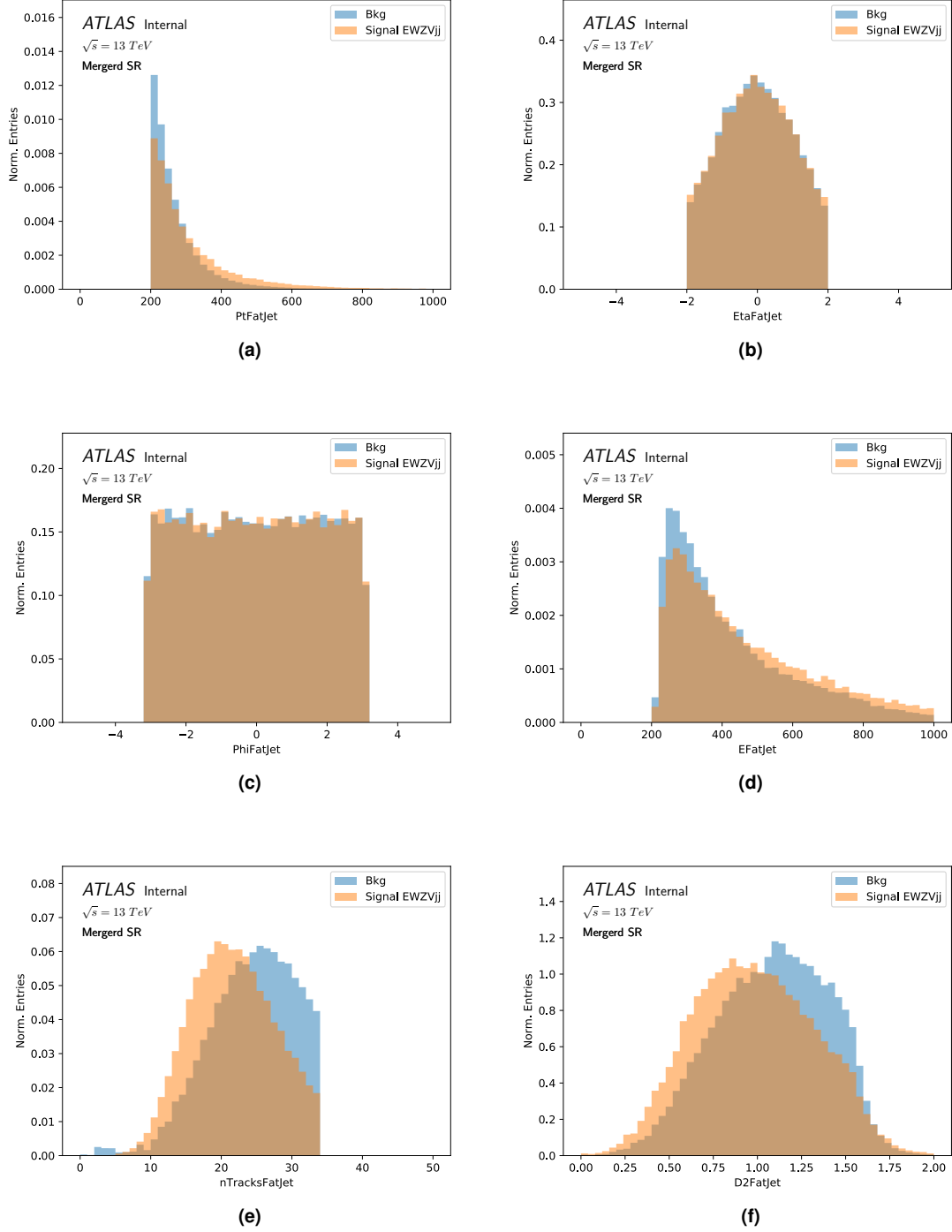


Figure 4.29: Distribution of the additional variables relative to the large- R jet implemented in the RNN training in the merged inclusive (HP + LP) signal region without the tagging jets selections, for signals (orange) and background (blue) events.

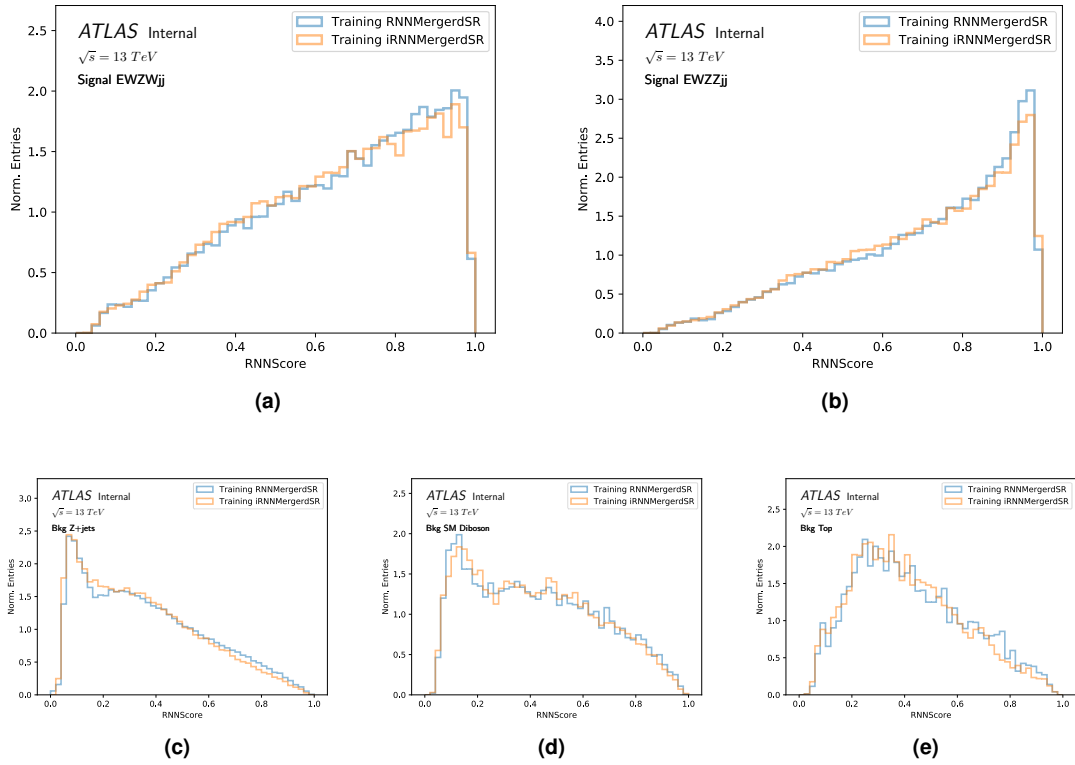


Figure 4.30: Distribution of the RNN score after the training in the merged inclusive (HP+SR) signal region (without the tagging jets selections) performed without (blue line) and with (orange line) the additional large-R jet variables as input features, for signals, Z+jets, diboson and top background events.

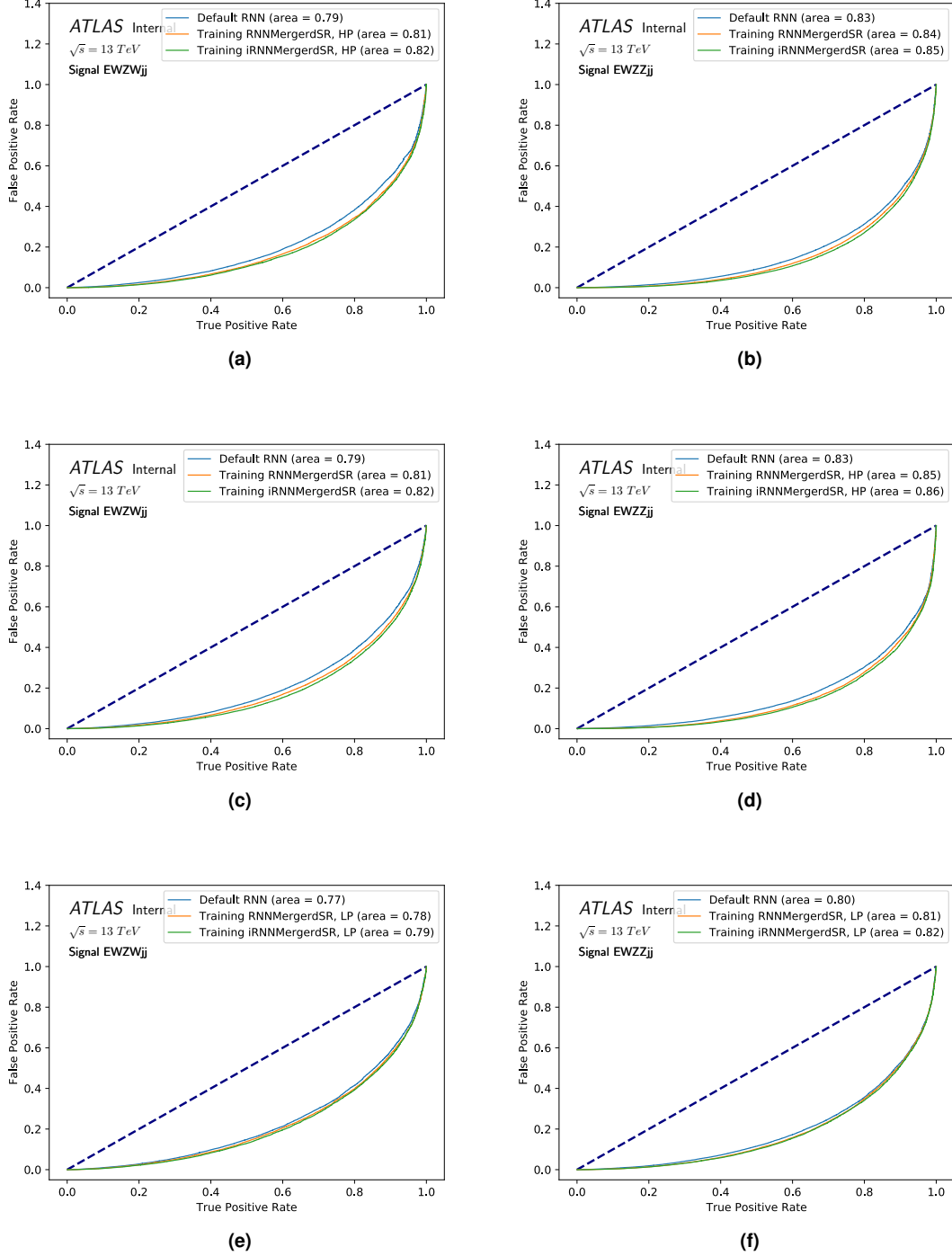


Figure 4.31: ROC curves after the dilepton training (azure line) and after the training in the merged inclusive (HP + LP) signal region (without the tagging jets selections) performed without (orange line) and with (green line) the additional large- R jets variables as input features, for ZW (first column) and ZZ (second column) signals. Both HP and LP signal regions events are considered in (a), (b), while only the HP signal region and only the LP signal region events are considered in (c), (d) and (e), (f), respectively.

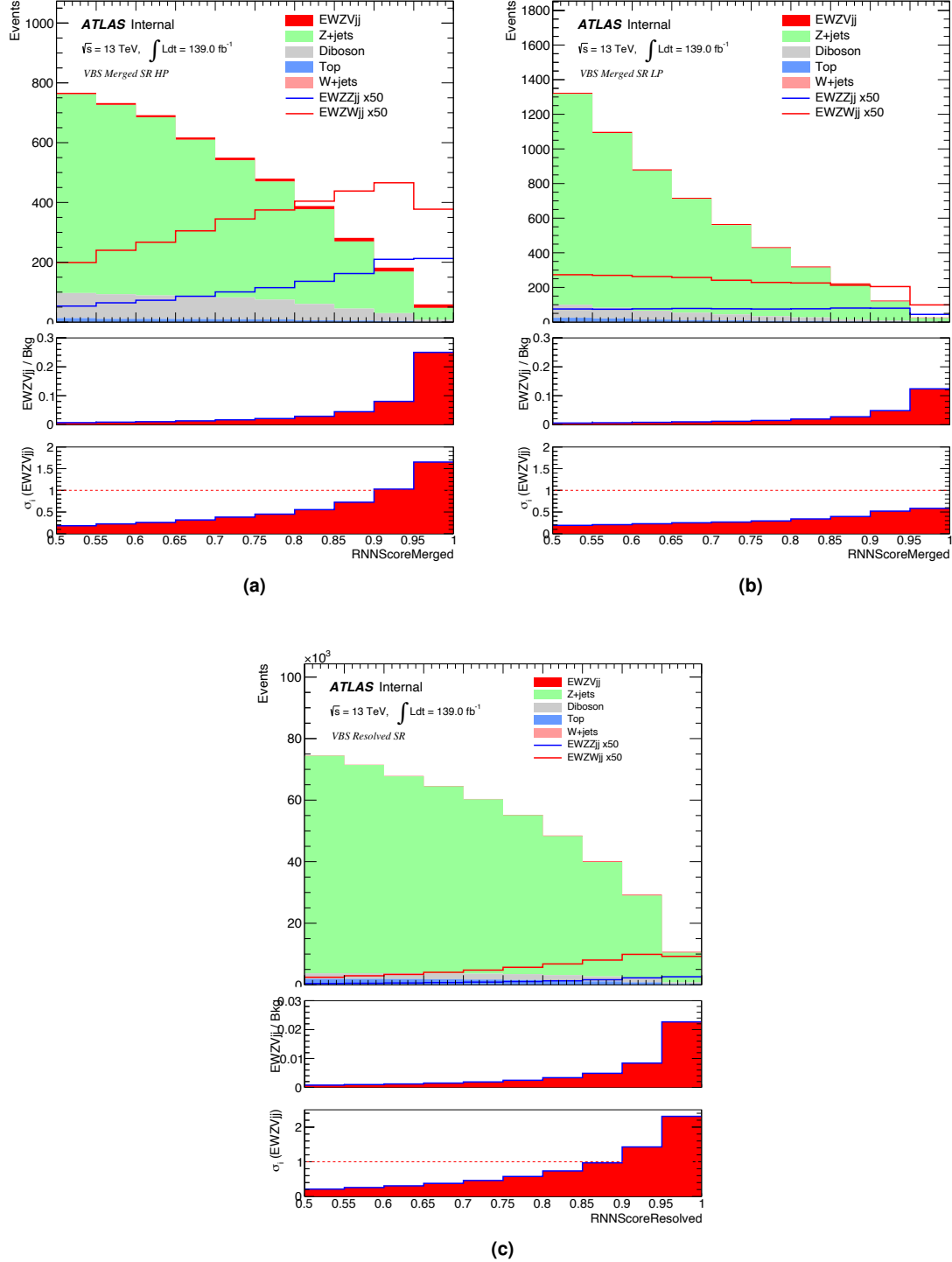


Figure 4.32: RNN score distribution for signal and background events in reference to the full 139 fb^{-1} integrated luminosity, in merged HP (a), merged LP (b) and resolved (c) signal region, when the RNN is trained in the signal regions (without the tagging jets selections). The two bottom pads display the signal-to-background ratio and the expected binned significance, evaluated on the RNN score.

4.4.3.1 RNN performance at high m_{VV}

The request for a high signal efficiency in the signal region is mainly driven by the idea of making aQGCs signal potentially detectable. As already mentioned, the effective field theory predicts an enhancement in the VBS cross section at high transverse momentum (p_T) of the vector bosons and at high values of the diboson system invariant mass (m_{VV}). Therefore, I have dealt with the check of the performance of the new RNN model (discussed in the previous section) for a m_{VV} window from 1 TeV up to 6 TeV, before and after applying some cuts (0.7, 0.8 and 0.9) to the RNN score. Actually, the simulated aQGCs samples were yet to be produced at the moment of this preliminary study, hence only the Standard Model VBS signal samples were exploited, bearing in mind that the aQGCs samples could only increase the yields in the high mass region. The full 139 fb^{-1} integrated luminosity was considered.

Before diving into the high mass region, a summary of the expected signal and background yields and efficiencies (the latter being evaluated with respect to the dilepton selection), as well as the binned significances, is given in Tables 4.20, 4.21, 4.22, for the merged HP, the merged LP and the resolved signal region, respectively. Both the cut-based values and the RNN values (before and after the score cuts) are taken into account. The expected binned significances are evaluated on m_{VV} . In the right side of the tables, the ratios between the RNN approach values and the cut-based values are displayed. When no cuts are applied to the score, both signal and background yields are enhanced with respect to the cut-based values, therefore the significance is a bit lower than the significance evaluated on the RNN score (Table 4.15). Nevertheless, even without applying any cut to the score, the significance for the ZW signal turns out to be of the same magnitude as its cut-based counterpart. On the other hand, applying the aforementioned cuts to the score, the significance results in being enhanced with respect to the cut-based approach. In reference to the 0.9 cut, the significance is about twice as high as the no-cut value. Figure 4.33 shows the m_{VV} full spectrum distribution in the signal regions, before and after the cuts are applied to the score.

Moving to the high mass region, the expected signal and background yields, efficiencies and significances as m_{VV} varies from 1 TeV to 6 TeV are displayed in figs. 4.34 to 4.42 for all the three signal regions. Since different mass bins are taken into account, the total significance 4.5 is illustrated instead of the asymptotic one.

For both signal and background events, the yields (and, consequently, the efficiencies) decrease as m_{VV} increases, as it is expected to happen. As regards the efficiencies and significances plots, the ratios between the RNN values and the cut-based values are displayed in the bottom pads of figs. 4.35, 4.36, 4.38, 4.39, 4.41 and 4.42. For all signal regions, the efficiency ratio is always above one for both ZZ and ZW signals, meaning that the signal efficiency is always enhanced with respect to the cut-based approach. In particular, the efficiency gain for the resolved signal region is sizeable, namely about 100%. Actually, the background yields are enhanced as well, though the significance does not seem to be affected: on the contrary, the significance values with respect to the cut-based counterparts are always above one, for all the signal regions and for both signals. Especially for the ZW signal, the significance enhancement is up to 100%.

The combined significance is estimated as the square root of the squaring sum of the significances obtained in the three signal regions separately. Table 4.17 shows the combined asymptotic significances in reference to the high mass region [1 TeV, 6 TeV] for the cut-based approach and the RNN approach, before and after applying the cuts to the score, as well as the RNN-CB ratios. On the other hand, Figure 4.43 displays the combined total significances as m_{VV} varies. When cutting the RNN score, the combined significance enhancement with respect to the cut-based case is up to 100%.

139 fb^{-1}						RNN $\div$ cut-based			
	Cut based	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9
$\sigma_{\text{asy}, ZZ} (m_{VV})$	5.62E-01	3.85E-01	6.48E-01	7.23E-01	8.27E-01	0.69	1.15	1.29	1.47
$\sigma_{\text{asy}, ZW} (m_{VV})$	1.36E+00	1.40E+00	2.23E+00	2.33E+00	2.37E+00	1.03	1.64	1.72	1.74

Table 4.17: Combined asymptotic significance (evaluated on m_{VV}) in the high mass region [1 TeV, 6 TeV], for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The right side of the table displays the ratios between the RNN values and the cut-based values.

In order to estimate the contribution of an individual signal region to the combined significance, a relative

significance is defined as

$$\sigma_{\text{rel}} = \frac{\sigma_i^2}{\sigma_{\text{combined}}^2} \quad (4.6)$$

where i may refer to the merged HP, the merged LP or the resolved signal region.

The relative asymptotic significances for the high m_{VV} region are given in table 4.18.

139 fb ⁻¹						RNN ÷ cut-based			
	Cut based	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9
<i>Merged HP SR</i>									
$\sigma_{\text{rel},ZZ}(m_{VV})$	6.00E-01	5.22E-01	4.90E-01	5.09E-01	5.41E-01	8.70E-01	8.16E-01	8.47E-01	9.01E-01
$\sigma_{\text{rel},ZW}(m_{VV})$	3.31E-01	3.26E-01	2.63E-01	2.59E-01	2.57E-01	9.85E-01	7.95E-01	7.82E-01	7.76E-01
<i>Merged LP SR</i>									
$\sigma_{\text{rel},ZZ}(m_{VV})$	8.00E-02	1.38E-01	1.19E-01	1.04E-01	7.81E-02	1.72E+00	1.48E+00	1.30E+00	9.77E-01
$\sigma_{\text{rel},ZW}(m_{VV})$	6.54E-02	1.12E-01	7.58E-02	6.62E-02	5.54E-02	1.71E+00	1.16E+00	1.01E+00	8.47E-01
<i>Resolved SR</i>									
$\sigma_{\text{rel},ZZ}(m_{VV})$	3.20E-01	3.40E-01	3.91E-01	3.88E-01	3.81E-01	1.06E+00	1.22E+00	1.21E+00	1.19E+00
$\sigma_{\text{rel},ZW}(m_{VV})$	6.04E-01	5.62E-01	6.61E-01	6.75E-01	6.88E-01	9.31E-01	1.09E+00	1.12E+00	1.14E+00

Table 4.18: Relative asymptotic significances (evaluated on m_{VV}) in the high mass region [1 TeV, 6 TeV], for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The right side of the table displays the ratios between the RNN values and the cut-based values.

Conversely, fig. 4.44 illustrates the relative total significances as m_{VV} varies. As it is made clear from the displayed plots, the estimated contributions of the individual signal regions to the combined significance in the high mass region [1 TeV, 6 TeV] can be approximated with the values reported in Table 4.19.

139 fb ⁻¹	Merged HP SR	Merged LP SR	Resolved SR
$\sigma_{\text{rel},ZZ/ZW}(\%)$	75	15	10

Table 4.19: Merged HP, merged LP and resolved signal region values of the relative significance in reference to the high mass region [1 TeV, 6 TeV]. The merged HP signal region provides the highest contribution to the combined significance.

The merged HP signal region seems to contribute to the combined significance for the most part, as regards the investigated high m_{VV} region. The merged LP and resolved signal regions provide the combined significance with comparable inputs.

In summation, not only does the new RNN training strategy (section 4.4.3) offer higher overall signal efficiencies and significances than the ones expected with the cut-based approach, it also yields a better performance at high values of the diboson mass system, especially in the merged HP signal region.

139 fb ⁻¹	Merged HP SR					Merged HP SR, RNN ÷ cutbased			
	Cut-based	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9
<i>Yields</i>									
VBS ZZ	1.38E+01	2.76E+01	1.87E+01	1.44E+01	8.45E+00	2.00	1.36	1.05	0.61
VBS ZW	2.53E+01	8.22E+01	4.81E+01	3.37E+01	1.69E+01	3.25	1.90	1.33	0.67
Bkg	1.20E+03	1.20E+04	1.87E+03	8.63E+02	2.16E+02	10.03	1.56	0.72	0.18
<i>Efficiency</i>									
VBS ZZ	2.63E-02	5.26E-02	3.57E-02	2.75E-02	1.61E-02	2.00	1.36	1.05	0.61
VBS ZW	1.04E-02	3.38E-02	1.98E-02	1.38E-02	6.92E-03	3.25	1.90	1.33	0.67
Bkg	7.90E-05	7.92E-04	1.24E-04	5.69E-05	1.42E-05	10.03	1.56	0.72	0.18
<i>$\sigma_{\text{asy}} m_{VV}$</i>									
VBS ZZ	4.35E-01	2.78E-01	4.54E-01	5.15E-01	6.08E-01	0.64	1.04	1.18	1.40
VBS ZW	7.81E-01	7.98E-01	1.14E+00	1.19E+00	1.20E+00	1.02	1.46	1.52	1.54

Table 4.20: Yields, efficiencies (with respect to the dilepton selection) and the binned significances (evaluated on m_{VV}) in the merged HP signal region, for the cut-based approach and the RNN approach, before and after the 0.7, 0.8 and 0.9 cuts are applied to the score. The right side of the table displays the ratio between the RNN values and the cut-based values.

139 fb ⁻¹	Merged LP SR					Merged LP SR, RNN ÷ cutbased			
	Cut-based	No rnn cut	Rnn score0.7	Rnn score0.8	Rnn score0.9	No rnn cut	Rnn score0.7	Rnn score0.8	Rnn score0.9
<i>Yields</i>									
VBS ZZ	8.48E+00	2.24E+01	8.66E+00	5.63E+00	2.48E+00	2.65	1.02	0.66	0.29
VBS ZW	1.91E+01	7.64E+01	2.42E+01	1.48E+01	6.08E+00	4.01	1.27	0.78	0.32
Bkg	3.19E+03	2.86E+04	1.66E+03	6.74E+02	1.41E+02	8.97	0.52	0.21	0.04
<i>Efficiency</i>									
VBS ZZ	1.62E-02	4.27E-02	1.65E-02	1.07E-02	4.73E-03	2.65	1.02	0.66	0.29
VBS ZW	7.83E-03	3.14E-02	9.95E-03	6.09E-03	2.49E-03	4.01	1.27	0.78	0.32
Bkg	2.11E-04	1.89E-03	1.10E-04	4.45E-05	9.33E-06	8.97	0.52	0.21	0.04
<i>$\sigma_{\text{asy}} m_{VV}$</i>									
VBS ZZ	1.59E-01	1.43E-01	2.23E-01	2.33E-01	2.31E-01	0.90	1.41	1.46	1.46
VBS ZW	3.47E-01	4.68E-01	6.13E-01	6.00E-01	5.57E-01	1.35	1.76	1.73	1.60

Table 4.21: Yields, efficiencies (with respect to the dilepton selection) and the binned significances (evaluated on m_{VV}) in the merged LP signal region, for the cut-based approach and the RNN approach, before and after the 0.7, 0.8 and 0.9 cuts are applied to the score. The right side of the table displays the ratio between the RNN values and the cut-based values.

139 fb ⁻¹	Resolved SR					Resolved SR, RNN ÷ cutbased			
	Cut-based	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9	No rnn cut	Rnn score>0.7	Rnn score>0.8	Rnn score>0.9
<i>Yields</i>									
VBS ZZ	1.38E+01	2.76E+01	1.87E+01	1.44E+01	8.45E+00	2.00	1.36	1.05	0.61
VBS ZW	2.53E+01	8.22E+01	4.81E+01	3.37E+01	1.69E+01	3.25	1.90	1.33	0.67
Bkg	1.20E+03	1.20E+04	1.87E+03	8.63E+02	2.16E+02	10.03	1.56	0.72	0.18
<i>Efficiency</i>									
VBS ZZ	2.63E-02	5.26E-02	3.57E-02	2.75E-02	1.61E-02	2.00	1.36	1.05	0.61
VBS ZW	1.04E-02	3.38E-02	1.98E-02	1.38E-02	6.92E-03	3.25	1.90	1.33	0.67
Bkg	7.90E-05	7.92E-04	1.24E-04	5.69E-05	1.42E-05	10.03	1.56	0.72	0.18
<i>$\sigma_{\text{asy}} m_{VV}$</i>									
VBS ZZ	4.35E-01	2.78E-01	4.54E-01	5.15E-01	6.08E-01	0.64	1.04	1.18	1.40
VBS ZW	7.81E-01	7.98E-01	1.14E+00	1.19E+00	1.20E+00	1.02	1.46	1.52	1.54

Table 4.22: Yields, efficiencies (with respect to the dilepton selection) and the binned significances (evaluated on m_{VV}) in the resolved signal region, for the cut-based approach and the RNN approach, before and after the 0.7, 0.8 and 0.9 cuts are applied to the score. The right side of the table displays the ratio between the RNN values and the cut-based values.

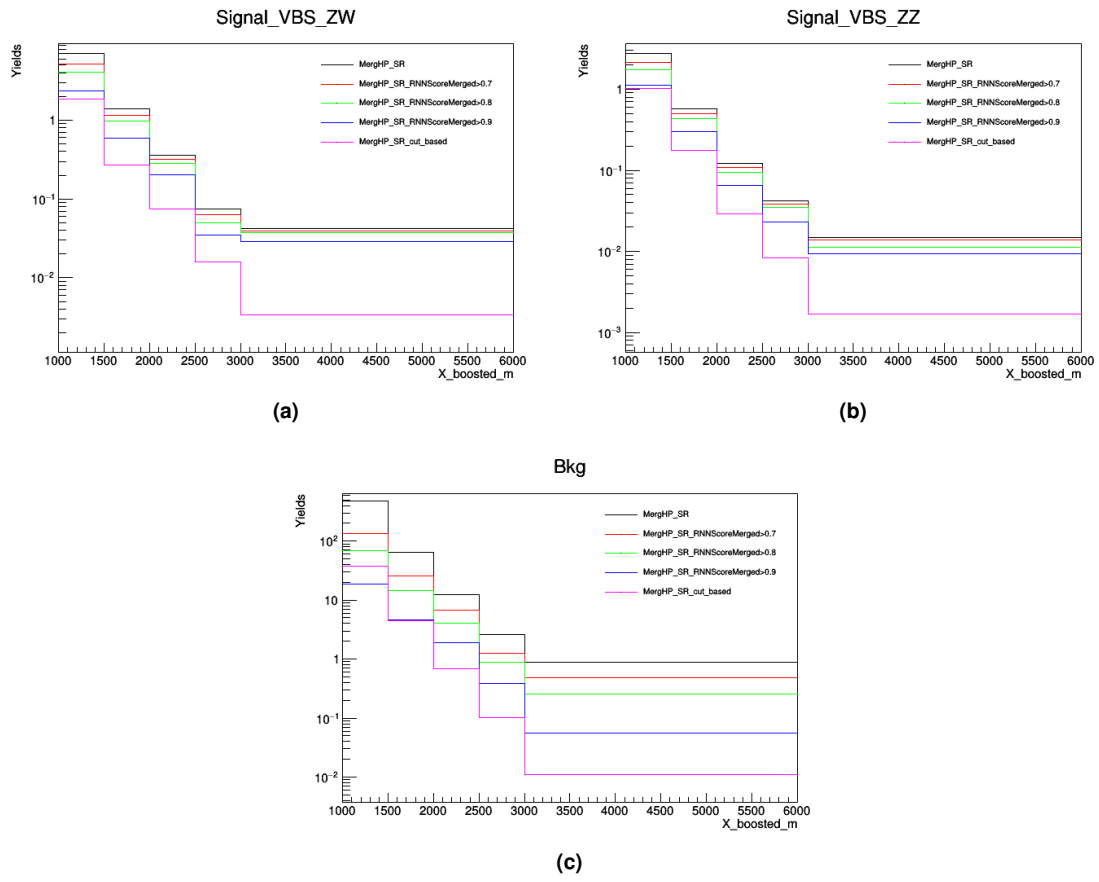


Figure 4.34: Signal and background yields as m_{VV} varies from 1 TeV to 6 TeV, in the *merged HP signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts.

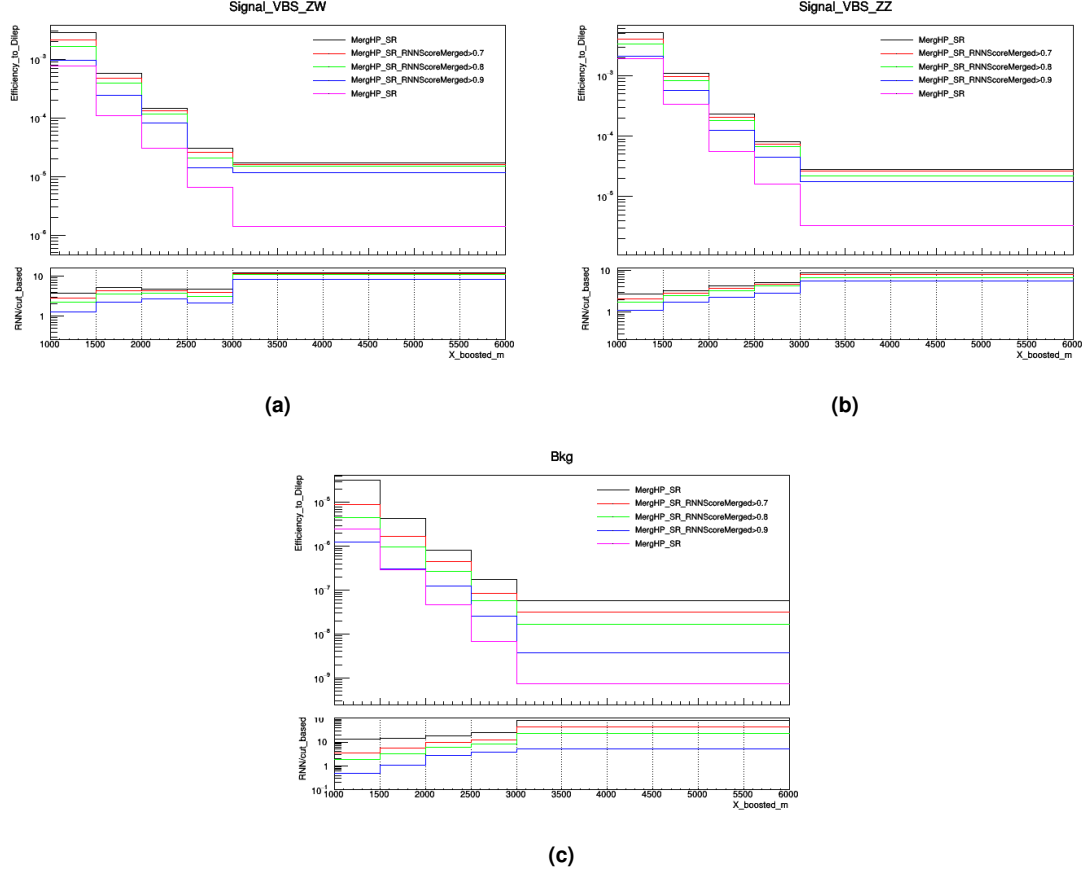


Figure 4.35: Signal and background efficiencies (evaluated with respect to dilepton selection) as m_{VV} varies from 1 TeV to 6 TeV, in the *merged HP signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The bottom pads display the ratios between the RNN values and the cut-based values.

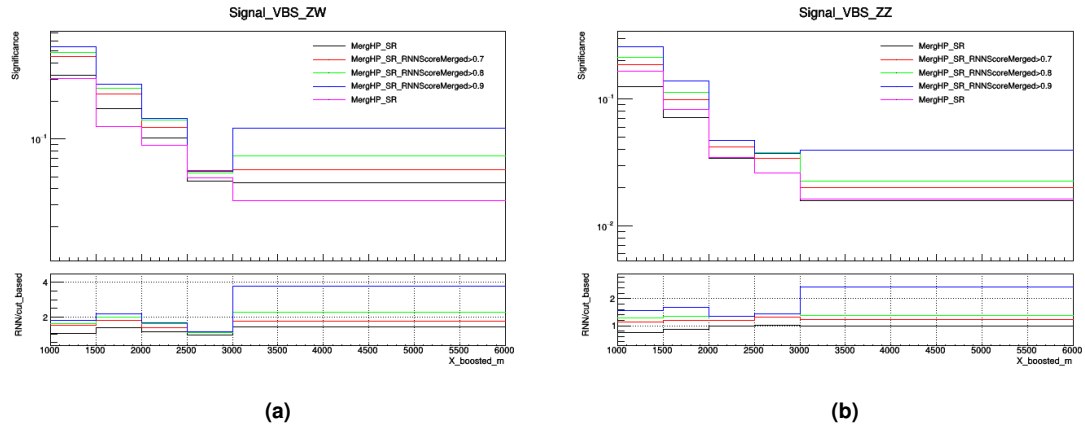


Figure 4.36: Total significances as m_{VV} varies from 1 TeV to 6 TeV, in the *merged HP signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The bottom pads display the ratios between the RNN values and the cut-based values.

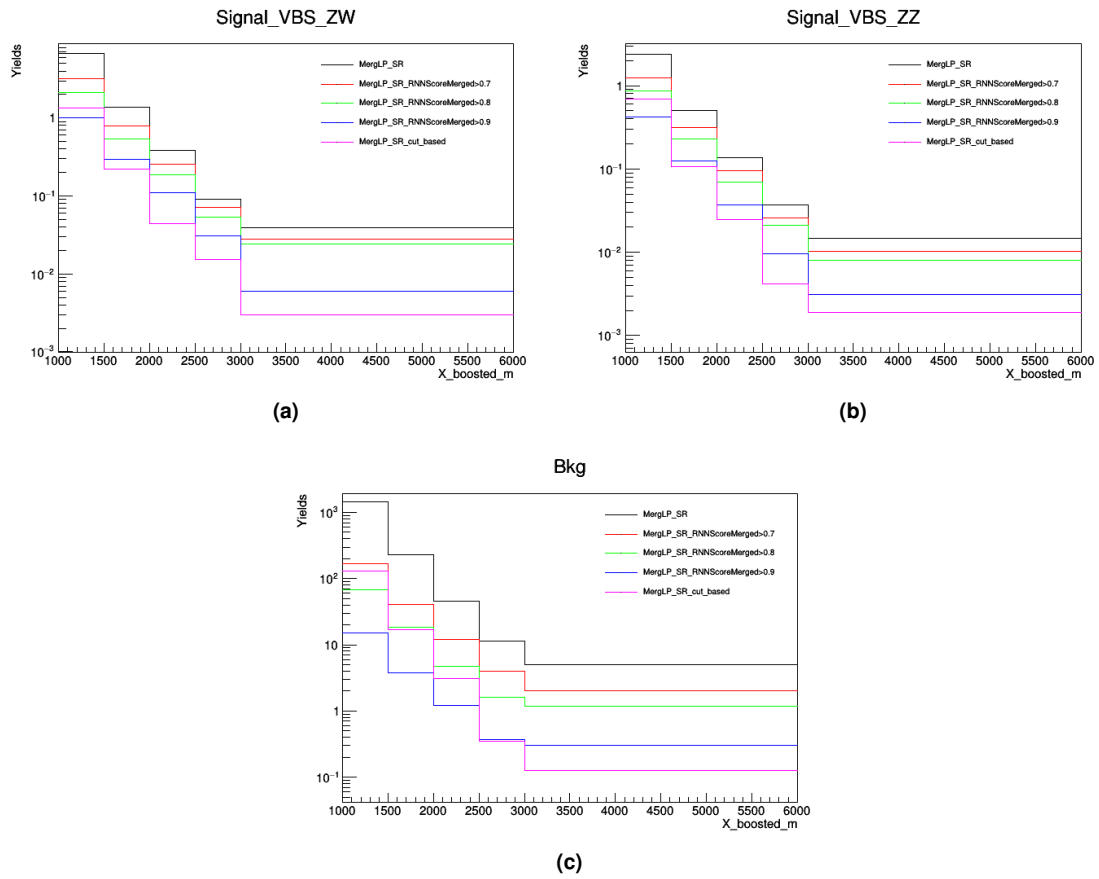


Figure 4.37: Signal and background yields as m_{VV} varies from 1 TeV to 6 TeV, in the *merged LP signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts.

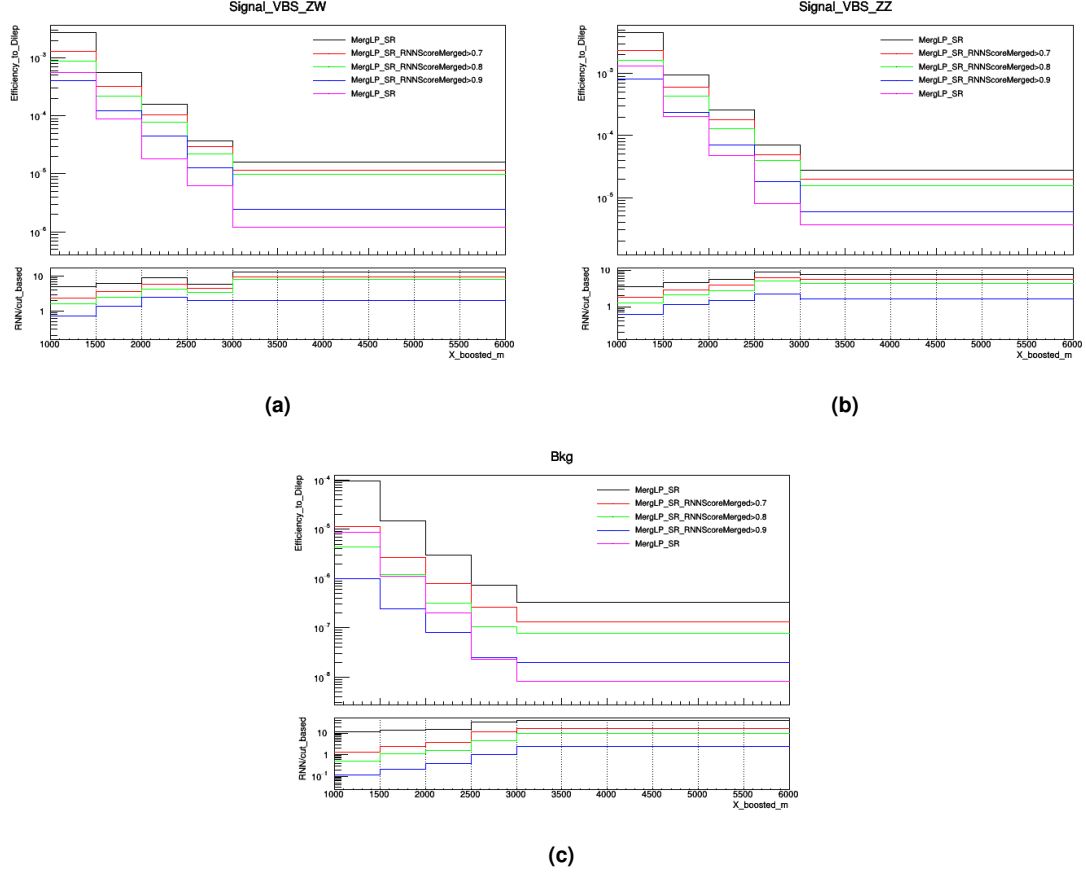


Figure 4.38: Signal and background efficiencies (evaluated with respect to the dilepton selection) as m_{VV} varies from 1 TeV to 6 TeV, in the *merged LP signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The bottom pads display the ratios between the RNN values and the cut-based values.

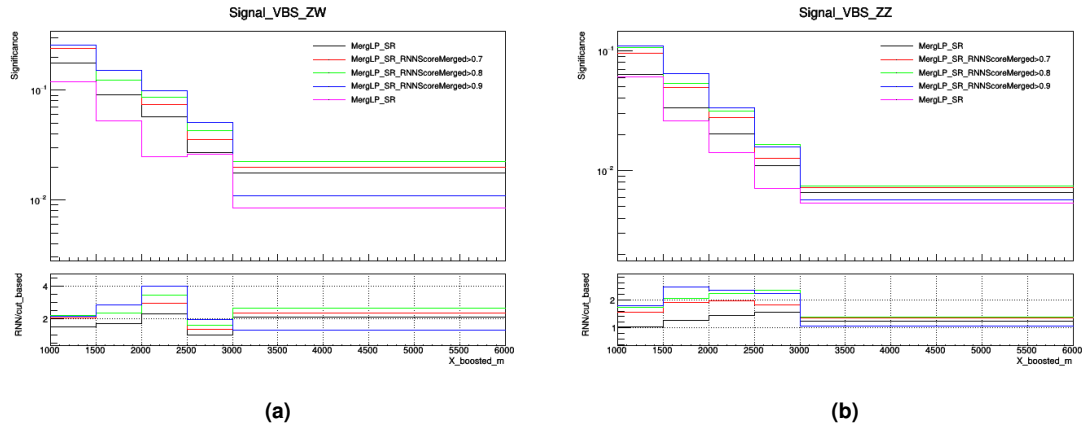


Figure 4.39: Total significances as m_{VV} varies from 1 TeV to 6 TeV, in the *merged LP signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The bottom pads display the ratios between the RNN values and the cut-based values.

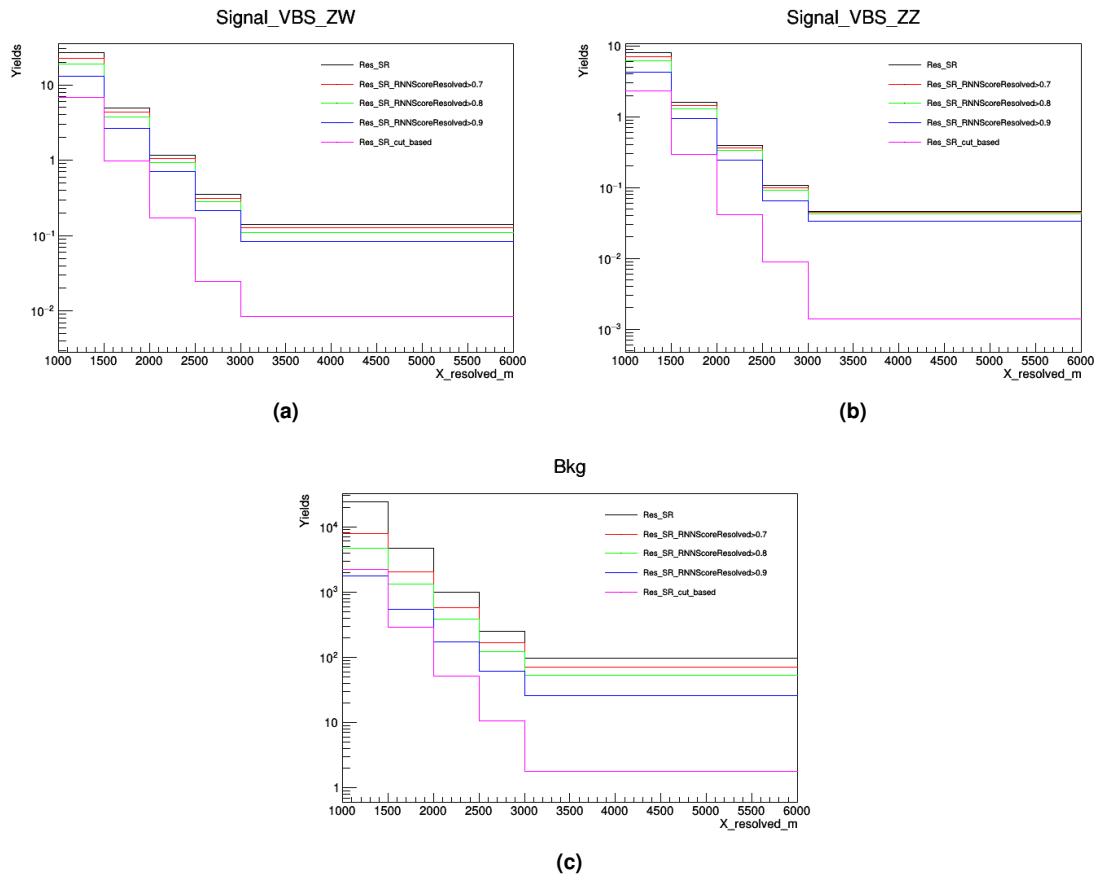


Figure 4.40: Signal and background yields as m_{VV} varies from 1 TeV to 6 TeV, in the *resolved signal region*, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts.

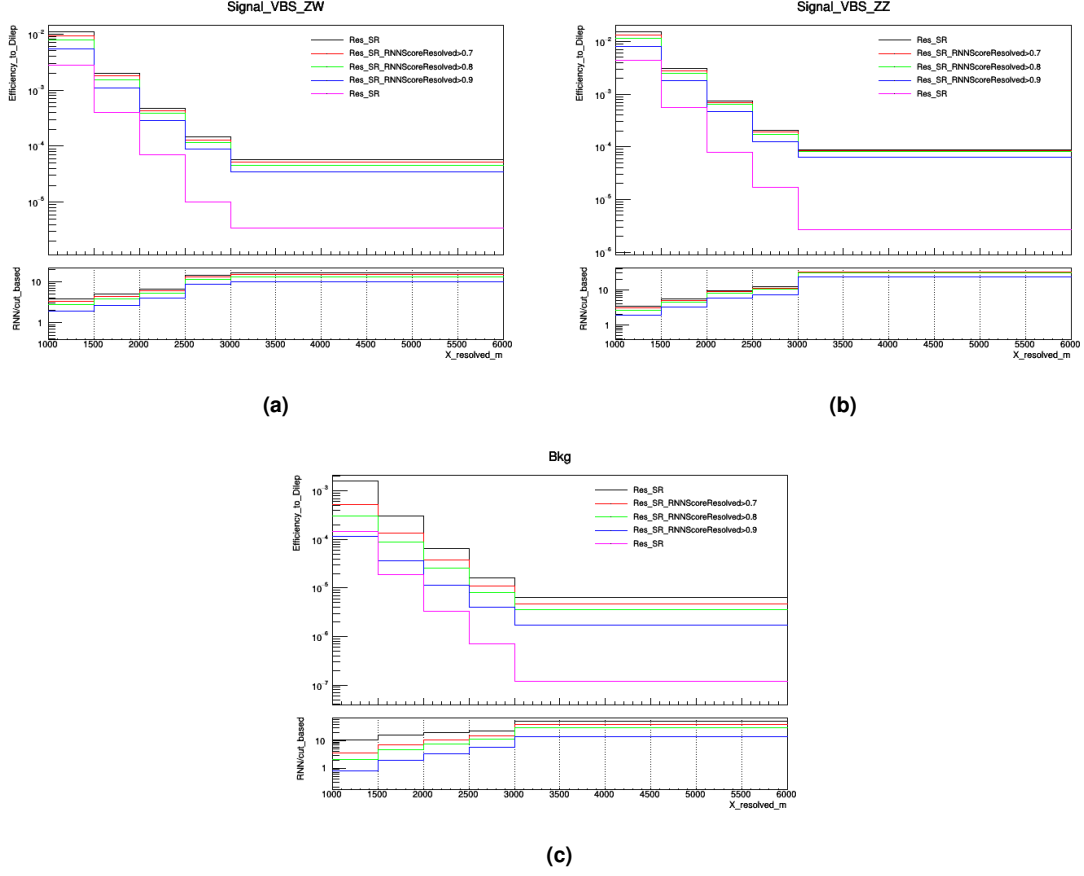


Figure 4.41: Signal and background efficiencies (evaluated with respect to the dilepton selection) as m_{VV} varies from 1 TeV to 6 TeV, in the **resolved signal region**, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The bottom pads display the ratios between the RNN values and the cut-based values.

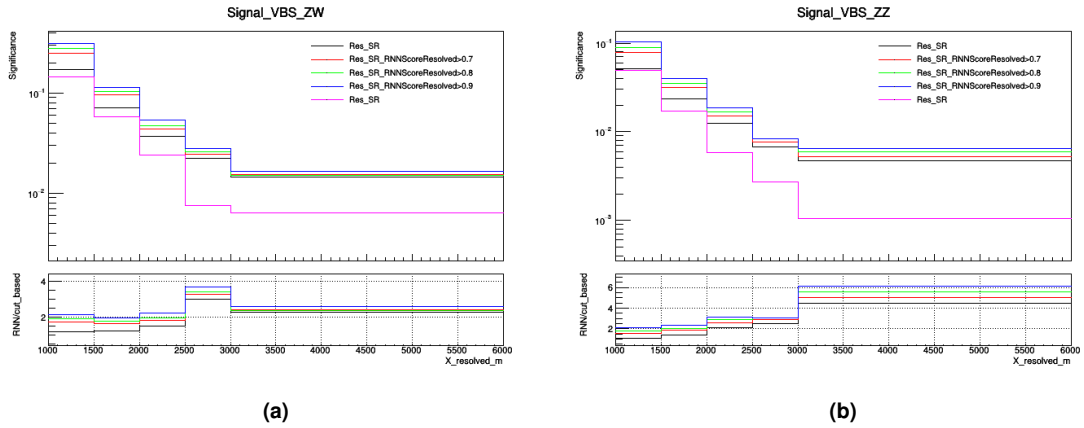


Figure 4.42: Total significances as m_{VV} varies from 1 TeV to 6 TeV, in the **resolved signal region**, for the cut-based approach and the RNN approach, before and after applying the 0.7, 0.8 and 0.9 score cuts. The bottom pads display the ratios between the RNN values and the cut-based values.

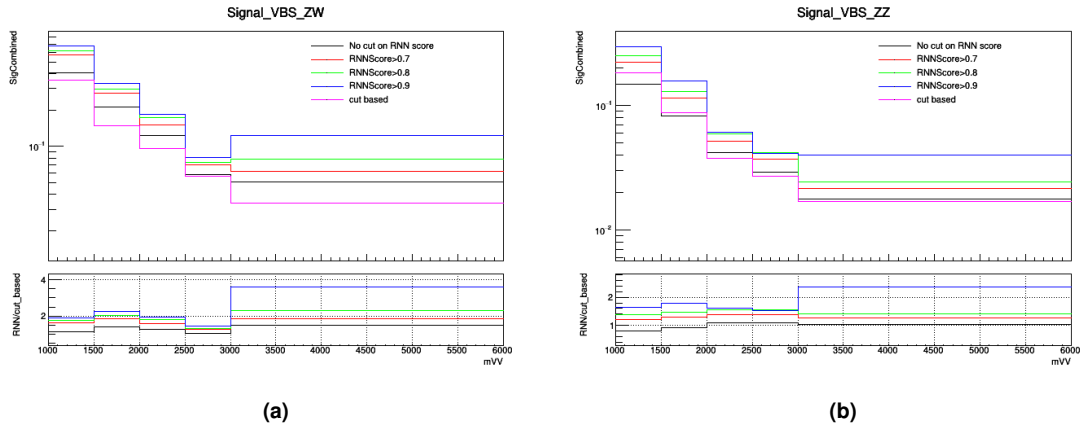


Figure 4.43: Combined significance as m_{VV} varies from 1 TeV to 6 TeV for the cut-based approach and the RNN approach, before and after the 0.7, 0.8 and 0.9 score cuts are applied. The bottom pads display the ratios between the RNN values and the cut-based values.

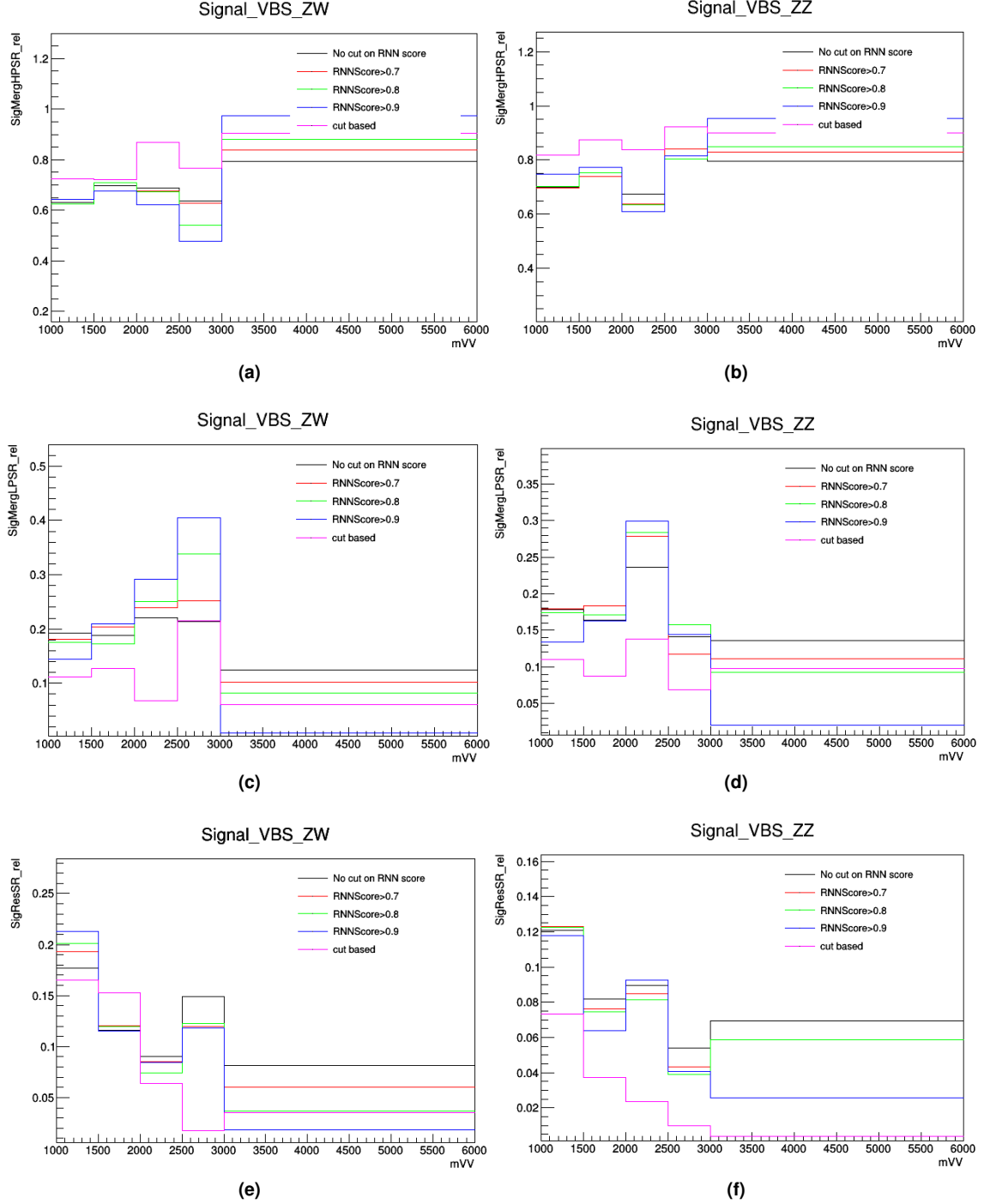


Figure 4.44: Relative ZW (first column) and ZZ (second column) signals significances in reference to the merged HP (first row), merged LP (second row) and resolved (third row) signal regions, as m_{VV} varies from 1 TeV to 6 TeV. The merged HP signal region seems to contribute for the most part to the combined significance.

4.4.3.2 Signal jets pairing: an alternative algorithm

As already explained in section 4.4.2.3, the signal jets pairing algorithm that designates the small- R jets pair whose invariant mass is the closest one to m_W or m_Z as the signal jets pair in the resolved selection, namely what has been referred to as the "close V" algorithm, has some bias issues related to the algorithm itself. An alternative criteria to select the signal jets pair would be selecting the two small- R jets with the highest p_T .

As regards the cut-based approach, the signal jets are searched among those jets that do not overlap with the tagging jets. As for the RNN approach, since the tagging jets are not defined, a different strategy has been pursued. In reference to the resolved case, the events with less than three small- R jets are discarded. With regard to the events with at least three small- R jets:

- if $N_{\text{jets}} > 3$ and a pair of jets with opposite-sign η is found, this very pair is removed from the list from which the signal jets are selected;
- if $N_{\text{jets}} > 3$ but a pair of jets with opposite-sign η is not found, or if $N_{\text{jets}} = 3$, the jet with the highest η (i.e., the most forward one) is removed from the list from which the signal jets are selected.

This removal procedure takes account of the fact that the first two leading jets are quite often tagging-like jets. However, the m_{jj}^{tag} cut is still avoided in order to retrieve signal efficiency. As for the merged regime, there are not any requests about the small- R jets: if at least one large- R jet is found, the one with the highest p_T is chosen as signal candidate.

The criteria upon which the signal jets are then selected (by means of what is referred to as the " p_T leading" algorithm) are summarized in Table 4.23.

	Merged SR	Resolved SR
	1 large- R jet	2 small- R jets
	leading p_T	leading and sub-leading p_T
Signal jets	50% (HP) or 80% (LP) WP boson tagging $D_2^{\beta=1}$, n_{track} and $m_{W/Z}$	$64 \text{ GeV} < m_{jj} < 106 \text{ GeV}$
	$p_T^I > 200 \text{ GeV}$, $ \eta_J < 2$	$p_T^{j_1} > 40 \text{ GeV}$, $p_T^{j_2} > 30 \text{ GeV}$
	$\Delta R(l_{1,2}, J) < 1.0$	

Table 4.23: Signal jets definition in the merged and resolved regime according to the p_T leading algorithm.

The RNN was re-trained in the signal regions defined by the p_T leading algorithm. The performance of the network in terms of the ROC is consistent with the previous training, which was performed in the signal regions defined by the close V algorithm (see section 4.4.3).

In order to validate the new model, I took care of investigating the agreement between the recorded data and the Monte Carlo prediction. Table 4.24 displays the full cut-flow signal, background and data yields, as well as the data-MC ratio and the asymptotic significances evaluated on the RNN score (without cutting the score) in the signal regions. The dilepton selection refers to the usual standard lepton requirements and the Z mass window plus the existence of at least two small- R jets. As for the pre-selection, it is defined just before the mass window requirements of the resolved signal region or just before the boson tagging requirements of the merged signal regions. The data-MC ratio is quite good through the full analysis flow, especially in the merged HP ZCR.

139 fb $^{-1}$	Dilepton	Merged Presel	Merged HP SR	Merged LP SR	Resolved Presel	Resolved SR	Merged HP ZCR	Merged LP ZCR	Resolved ZCR
VBS ZZ	5.20E+02	9.40E+01	2.80E+01	2.20E+01	4.60E+02	1.30E+02	5.00E+00	6.70E+00	2.50E+02
VBS ZW	2.40E+03	3.30E+02	8.20E+01	7.60E+01	2.20E+03	6.60E+02	2.00E+01	3.00E+01	1.10E+03
Bkg	6.20E+06	1.30E+05	1.20E+04	2.90E+04	5.30E+06	4.50E+05	6.50E+03	1.30E+04	1.30E+06
Data	6.00E+06	1.20E+05			5.20E+06		6.50E+03	1.20E+04	1.20E+06
Data/MC	0.973	0.922			0.987		0.995	0.916	0.932
$\sigma_{\text{asy}, ZZ}(\text{RNNScore})$			0.65	0.29		0.73			
$\sigma_{\text{asy}, ZW}(\text{RNNScore})$			1.46	0.82		2.33			

Table 4.24: Full cut-flow yields, data-MC ratio and asymptotic significances (evaluated on the score, without any score cut) as regards the RNN training in the signal regions defined by the p_T leading algorithm.

Table 4.25 shows the signal and background efficiencies with respect to the dilepton selection (when at least two small- R jets are found), both for the RNN and the cut-based approach, as well as the RNN-CB efficiency ratios. The signal efficiency enhancement is up to a factor 4, depending on the signal.

	Merged Presel	Merged HP SR	Merged LP SR	Resolved Presel	Resolved SR
<i>RNN</i>					
VBS ZZ	0.183	0.053	0.043	0.883	0.249
VBS ZW	0.138	0.035	0.032	0.907	0.276
Bkg	0.020	0.002	0.005	0.852	0.073
<i>Cut-based</i>					
VBS ZZ	0.061	0.027	0.016	0.363	0.141
VBS ZW	0.031	0.011	0.008	0.274	0.100
Bkg	0.002	0.0002	0.001	0.055	0.011
<i>RNN $\div$ cut-based</i>					
VBS ZZ	3.0	2.0	2.6	2.4	1.8
VBS ZW	4.5	3.3	4.0	3.3	2.8

Table 4.25: Signal and background efficiencies with respect to the dilepton selection (when at least two small- R jets are found), for the RNN approach and the cut-based approach. The RNN-CB efficiency ratios are displayed as well.

Following the analysis flow, dilepton selection is the starting point. The lepton variables are taken into considerations in Figures 4.45 and 4.46: the global data-MC ratio is about 97% and the overall agreement is quite good as well, as it is made clear from the ratio plots. As for the merged pre-selection, the variables regarding the signal candidate large- R jet are displayed (Figure 4.47), such as the four-momentum variables as well as the number of tracks, the $D_2^{\beta=1}$, the large- R jets multiplicity and the invariant mass m_J . The global data-MC agreement is closer to unity in the case of the resolved pre-selection. Here (Figure 4.48) the signal candidate small- R jets four-momentum is illustrated alongside with the small- R jets multiplicity and the invariant mass of the signal candidate pair.

Moving to the control regions, the four-momentum variables and the number of tracks of the first three small- R jets are taken into account (Figures 4.50, 4.52, 4.56), since the RNN training is performed in the signal regions employing these very variables for the first six jets (referring to the merged case, the $D_2^{\beta=1}$ is deployed as well). The modelling of the input features turns out to be notable, especially in the merged regime. As for the resolved selection, the ratio plots for the first three jets p_T is not as flat as in the merged case (Figures 4.56a, 4.56f, 4.56k), however the number of the events involved with this mismodelling is a small fraction of the total events.

For both the merged high and the low purity control regions, the data-MC ratio plots for the RNN score and the diboson system invariant mass m_{VV} are remarkably flat and close to unity (Figures 4.53, 4.54), even though the global data-MC agreement in the low purity region is a bit lower than in the high purity region (due to the $D_2^{\beta=1}$ modelling). On the other hand, the resolved case seems to suffer from mismodelling issues for the same variables (Figure 4.57). This is probably related to the Monte Carlo generator employed for the Z +jets background modelling, namely SHERPA (section 4.1.2). Implementing the RNN model trained on MC samples with Z +jets generated by SHERPA and evaluating the score on MC samples with Z +jets generated by MADGRAPH results in a better data-MC agreement, as it is made clear from Figure 4.58a. As a matter of fact, the mismodelling does not seem to depend on the RNN architecture: training the network using the MADGRAPH samples for Z +jets, the score is not affected by the different training (Figure 4.58b).

In conclusion, the general background modelling is quite satisfying. In particular, the RNN architecture trained in the signal regions defined by the " p_T leading" algorithm provides a compelling data-MC agreement in the control regions when it comes to the RNN input variables, the RNN score and the invariant mass of the diboson system.

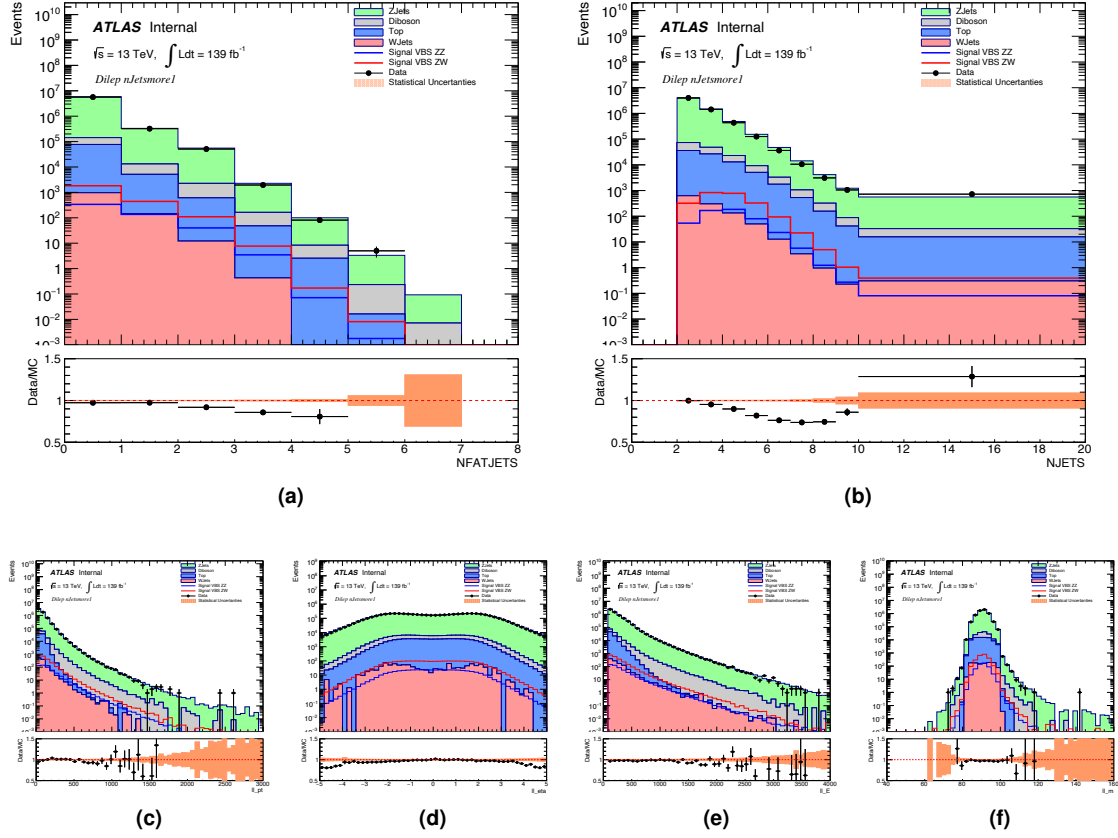


Figure 4.45: Distribution of the large- R jets and small- R jets multiplicity (a), (b), and the four-momentum variables of the dilepton system (c) - (f), at dilepton level (when at least two small- R jets are found). The bottom pads display the data-MC ratio.

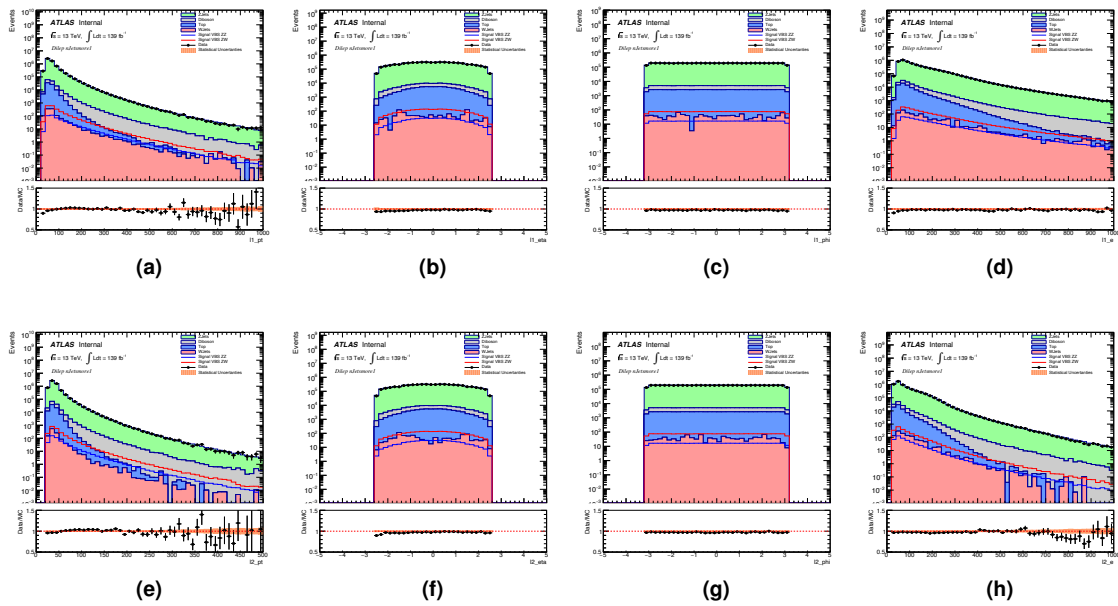


Figure 4.46: Four-momentum variables of the leading (first row) and sub-leading (second row) lepton of the dilepton system, at dilepton level (when at least two small- R jets are found). The bottom pads display the data-MC ratio.

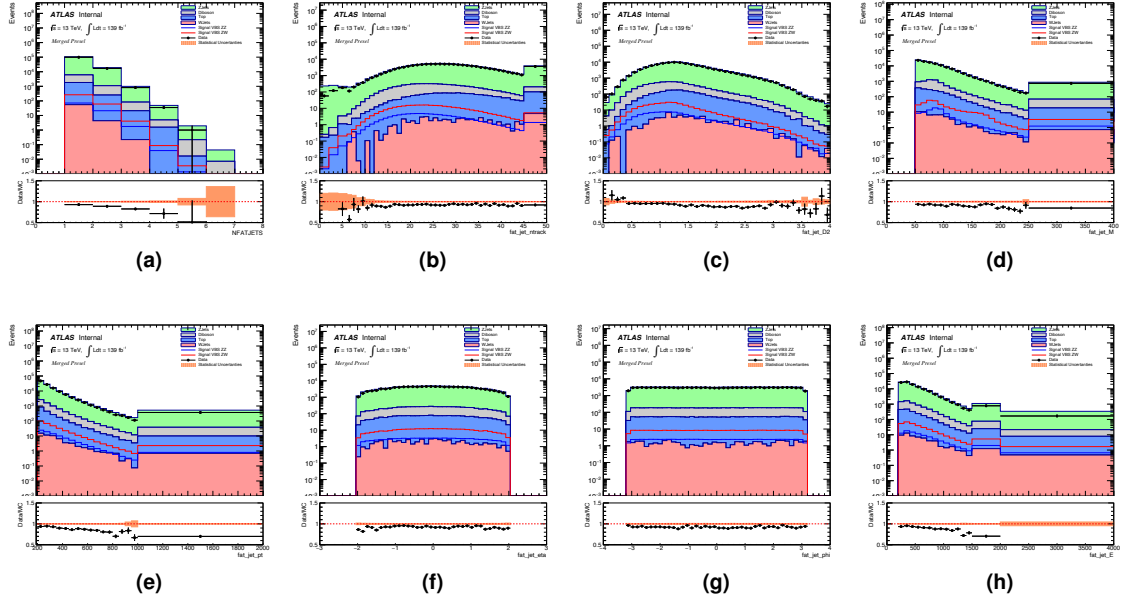


Figure 4.47: First row: distribution of the large- R jets multiplicity, the number of tracks, $D_2^{\beta=1}$, m_J of the signal candidate large- R jet, in the merged pre-selection. Second row: distribution of the four-momentum variables of the signal candidate large- R jet, in the merged pre-selection. The bottom pads display the data-MC ratio.

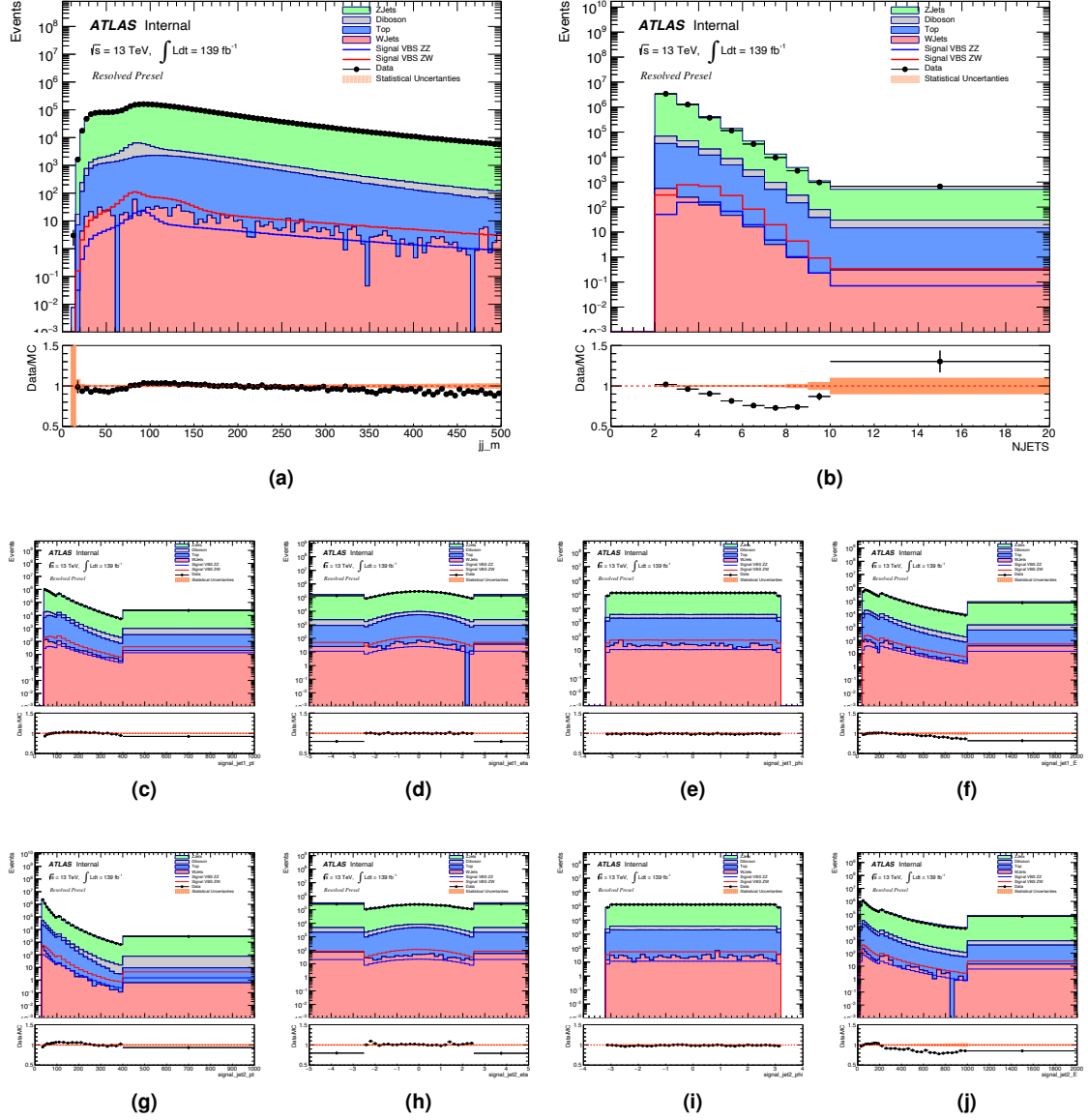


Figure 4.48: First row: distribution of the signal jets pair invariant mass and small- R jets multiplicity. Second and third row: four-momentum variables of the leading and sub-leading signal jets, respectively. All distribution are evaluated in the resolved pre-selection. The bottom pads display the data-MC ratio.

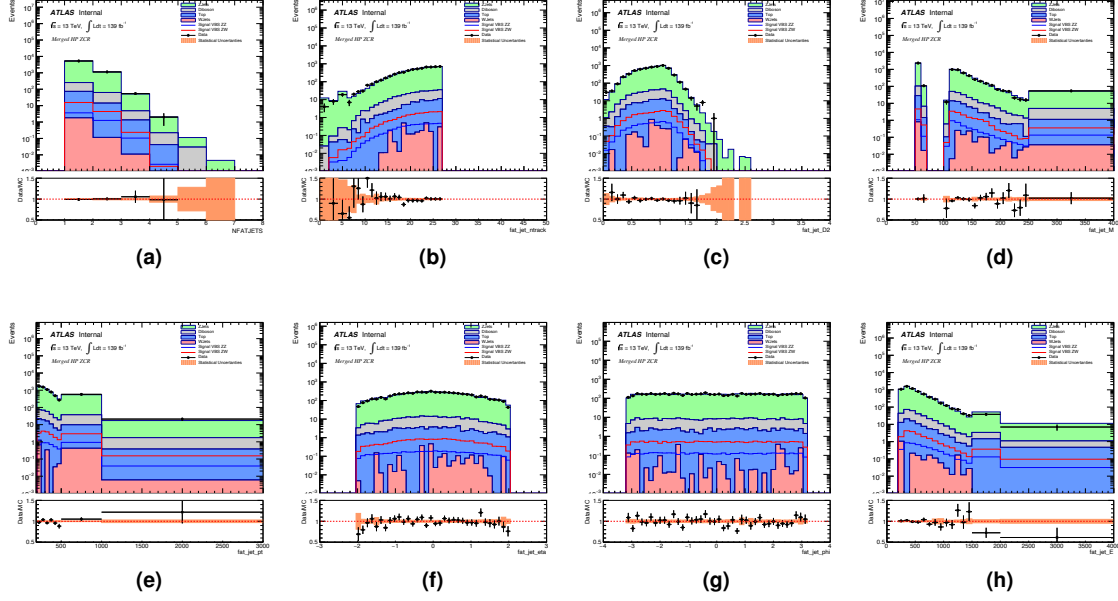


Figure 4.49: First row: distribution of the large- R jets multiplicity, the number of tracks, $D_2^{\beta=1}$, m_J of the signal large- R jet, in the merged high purity ZCR. Second row: distribution of the four-momentum variables of the signal large- R jet, in the merged high purity ZCR. The bottom pads display the data-MC ratio.

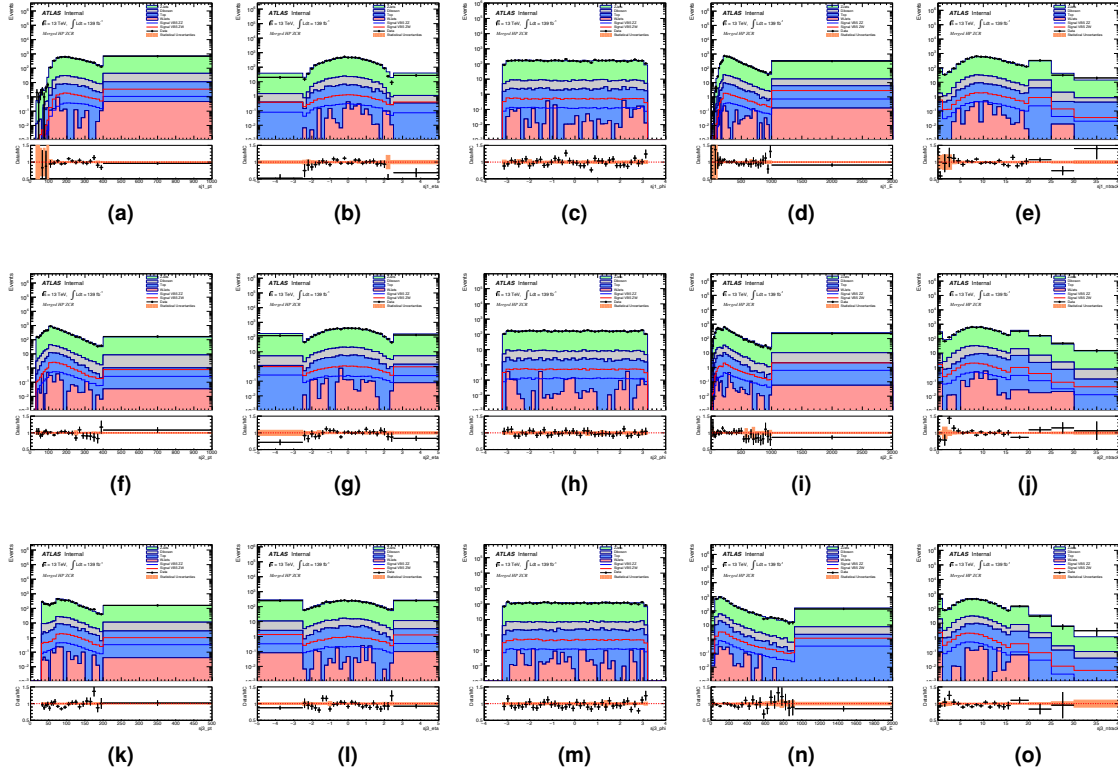


Figure 4.50: Distribution of the four-momentum variables, plus the number of tracks, of the leading (first row), sub-leading (second row) and sub-sub-leading (third row) small- R jets, in the merged high purity ZCR. The bottom pads display the data-MC ratio.

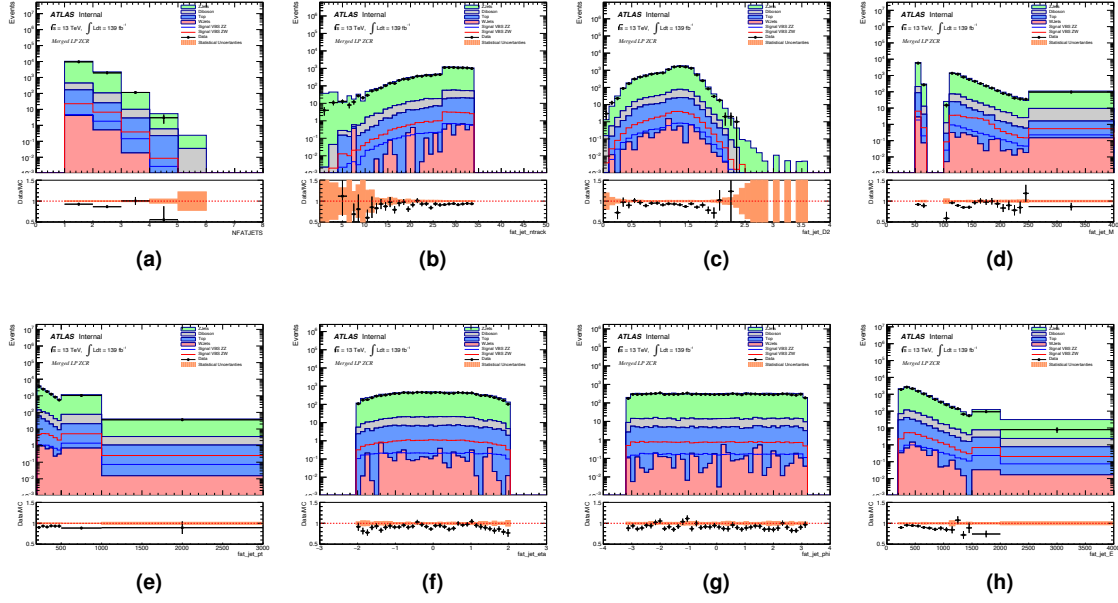


Figure 4.51: First row: distribution of the large- R jets multiplicity, the number of tracks, $D_2^{\beta=1}$, m_J of the signal large- R jet, in the merged low purity ZCR. Second row: distribution of the four-momentum variables of the signal large- R jet, in the merged low purity ZCR. The bottom pads display the data-MC ratio.

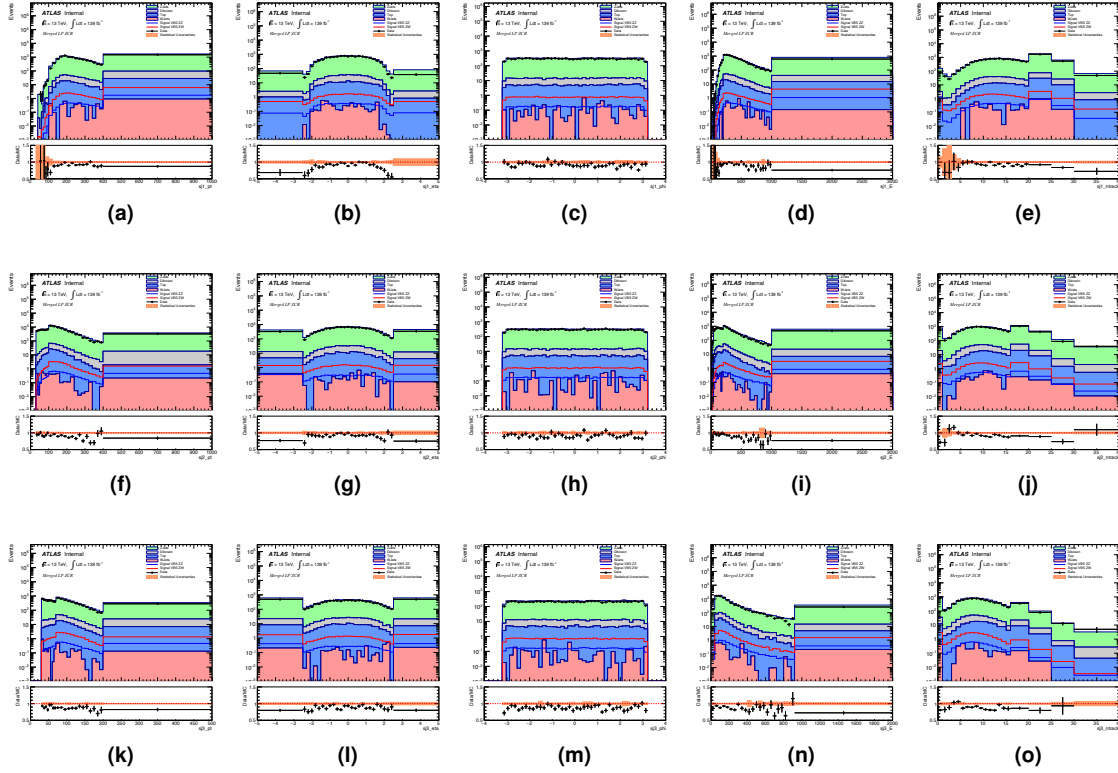


Figure 4.52: Distribution of the four-momentum variables, plus the number of tracks, of the leading (first row), sub-leading (second row) and sub-sub-leading (third row) small- R jets, in the merged low purity ZCR. The bottom pads display the data-MC ratio.

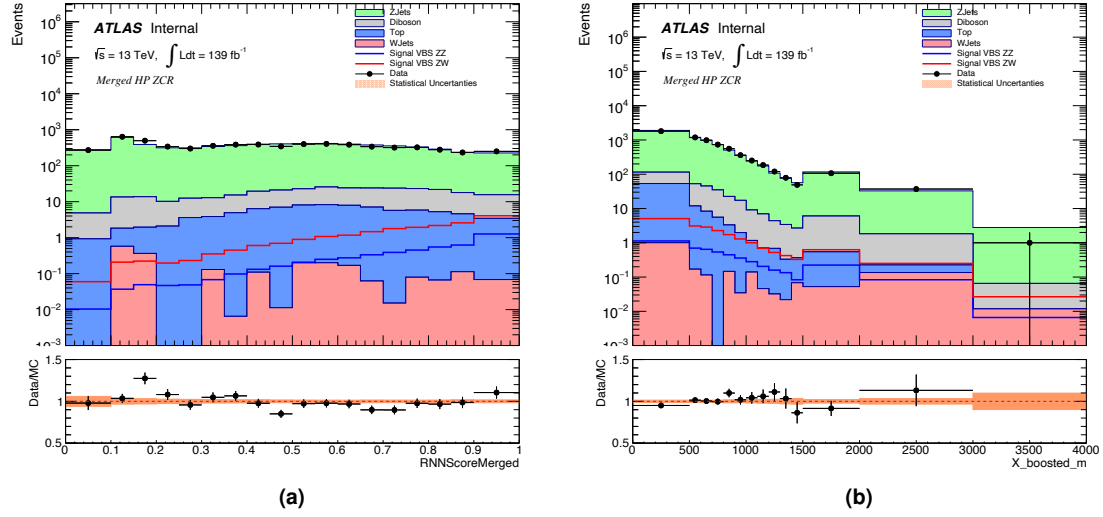


Figure 4.53: Distribution of the RNN score and the diboson system invariant mass in the merged high purity ZCR. The bottom pads display the data-MC ratio.

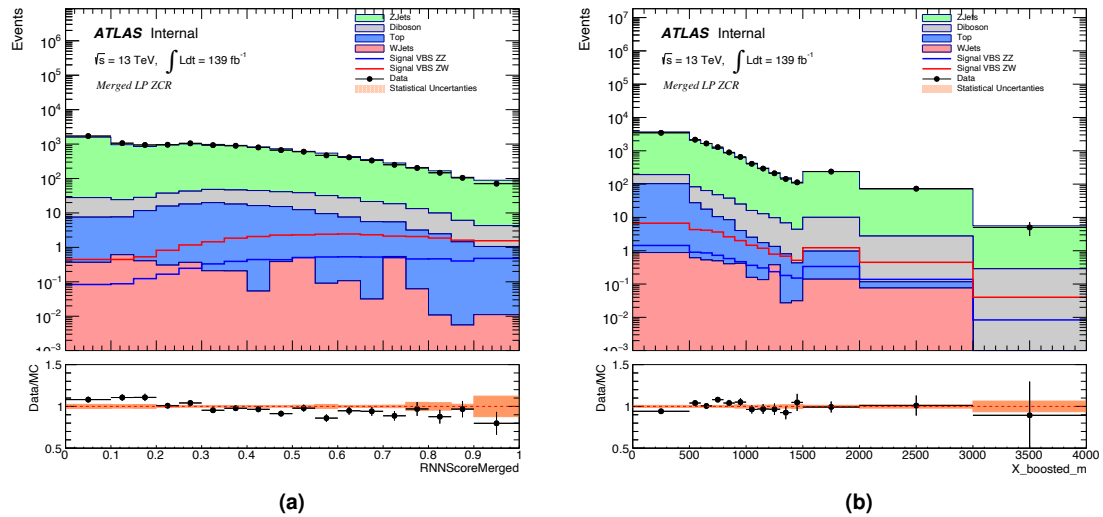


Figure 4.54: Distribution of the RNN score and the diboson system invariant mass in the merged low purity ZCR. The bottom pads display the data-MC ratio.

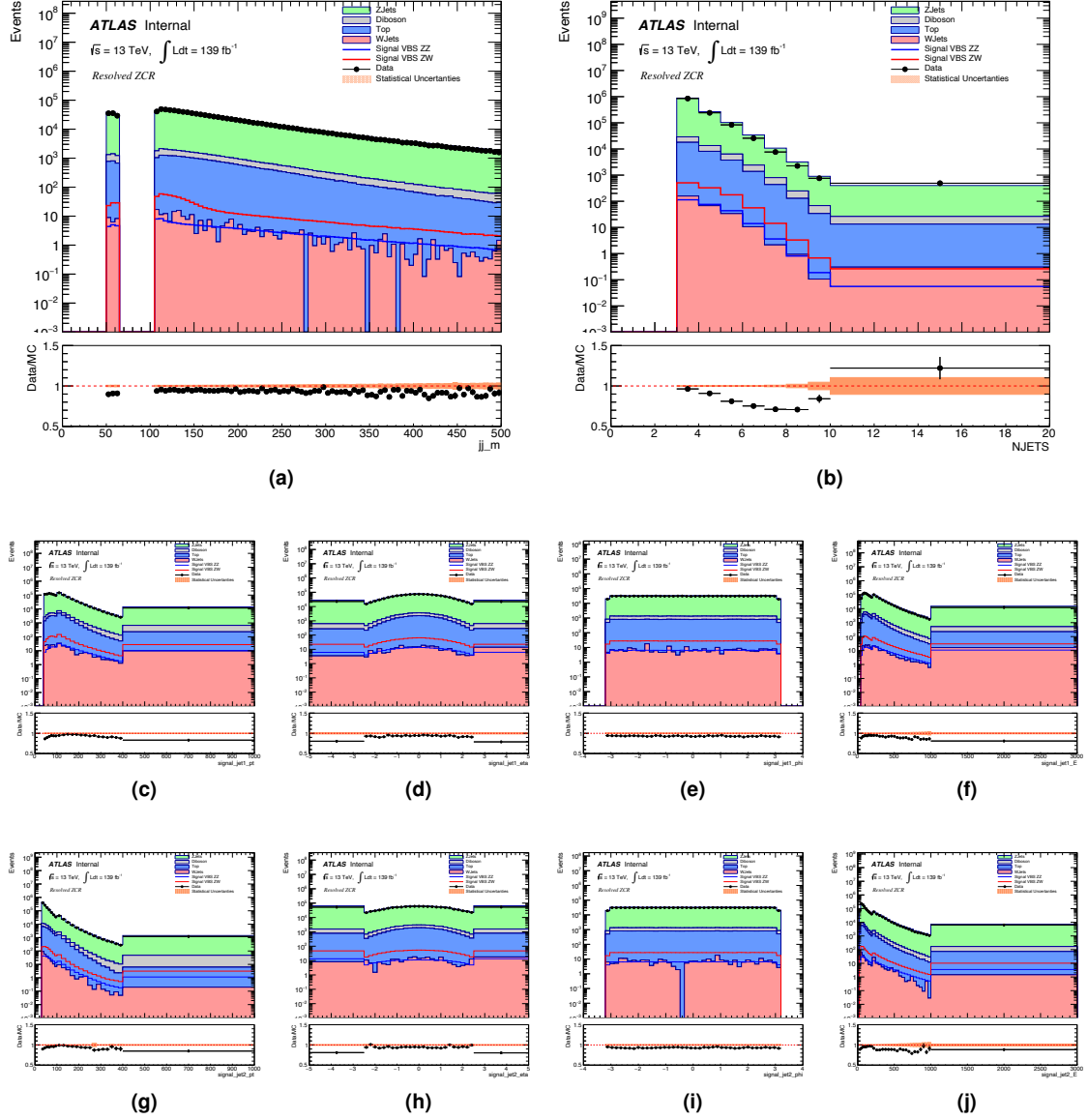


Figure 4.55: First row: distribution of the signal jets pair invariant mass and small- R jets multiplicity. Second and third row: four-momentum variables of the leading and sub-leading signal jets, respectively. All distribution are evaluated in the resolved ZCR. The bottom pads display the data-MC ratio.

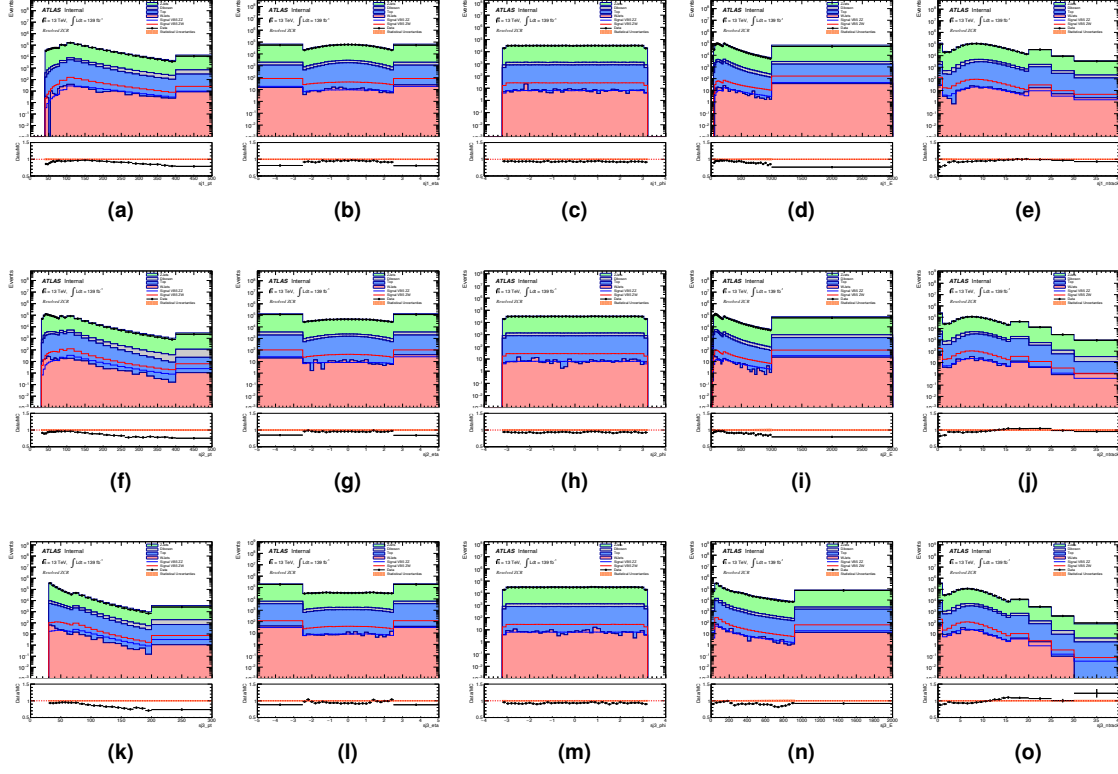


Figure 4.56: Distribution of the four-momentum variables, plus the number of tracks, of the leading (first row), sub-leading (second row) and sub-sub-leading (third row) small- R jets, in the resolved ZCR. The bottom pads display the data-MC ratio.

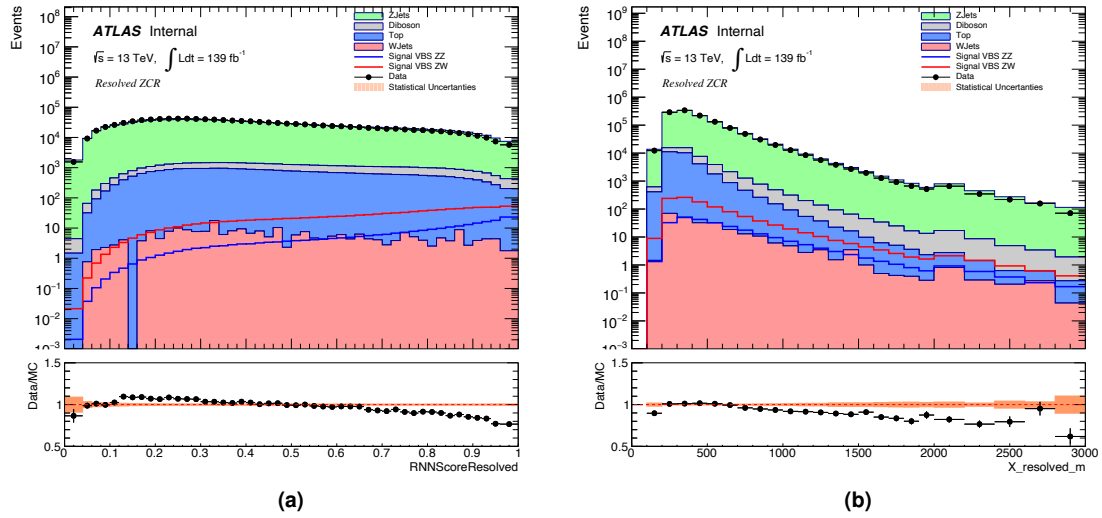


Figure 4.57: Distribution of the RNN score and the diboson system invariant mass in the resolved ZCR.

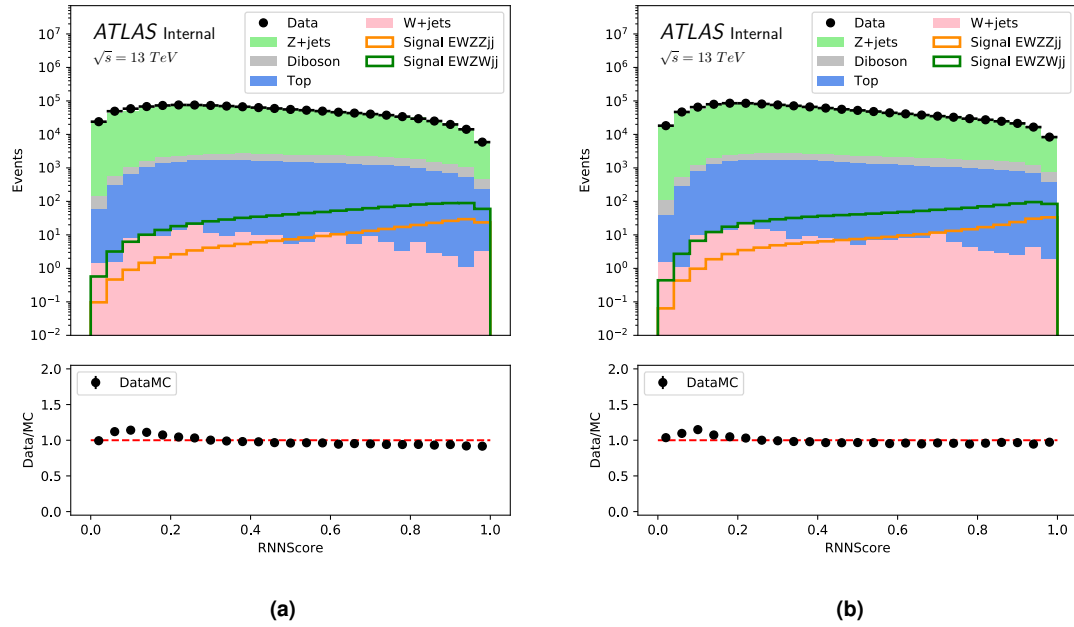


Figure 4.58: RNN score distribution in the resolved ZCR with Z+jets background generated with MADGRAPH, training the RNN with SHERPA Z+jets samples (a) and with MADGRAPH Z+jets samples (b). The bottom pads display the data-MC ratio, which is not affected by the RNN training.

Chapter 5

Statistical analysis

The current search for electroweak diboson production in semileptonic final states performed by the ATLAS Collaboration illustrated in the previous chapter is, at present, at a very preliminary stage. In fact, as already pointed out, my Thesis work has been mainly devoted to the optimisation of the current event selection strategy, by means of the employment of a Recurrent Neural Network. The analysis has still a long way to go before the definition of a final discriminant to utilise in order to perform a cross-section measurement. Similarly, EFT Monte Carlo samples are still into production, so that measurements regarding aQGCs are long-term goals as well.

Nonetheless, the general statistical strategy is already designed by the previous run search approach, which will be summarised in the following section.

5.1 Previous run statistical approach

In the previous run search [31], a multivariate method was employed in order to enhance the separation between the signal and the background and define a final discriminant to whose distribution perform a fit and extract the cross-section measurement. The Toolkit for Multivariate Data Analysis, TMVA [1], was used with the Boosted Decision Trees (BDTs) methods. The BDT training was performed using simulated signal and all background Monte Carlo samples. For each channel a minimal set of input variables which maximizes the BDT discriminant separation between signal and background was selected.

The BDT outputs were used alongside with other variables distributions to perform a combined profile likelihood fit by means of the statistical framework `WMaker` (developed by the ATLAS Collaboration), where the signal strength of the SM VBS (WW/WZ/ZZ) processes was assigned as the POI (parameter of interest) and the systematic uncertainties together with the background normalization factors ($W + \text{jets}$, $Z + \text{jets}$, $t\bar{t}$) were assigned as NPs (Nuisance parameters).

5.1.1 Likelihood function and statistical hypothesis test

The statistical analysis of the data used a binned likelihood function constructed as

$$\mathcal{L}(\mu, \theta) = \prod_{b=1}^{N_{\text{bins}}} \frac{(\mu\nu_b^{\text{sig}} + \nu_b^{\text{bkg}})^{n_b}}{n_b!} e^{-(\mu\nu_b^{\text{sig}} + \nu_b^{\text{bkg}})} \quad (5.1)$$

where:

- μ is the signal strength parameter to extract from the fit;
- N_{bins} is the number of bins of the input histogram;
- ν_b^{sig} and ν_b^{bkg} are the expected signal and background yields, respectively, in each histogram bin, so that $\mu\nu_b^{\text{sig}} + \nu_b^{\text{bkg}}$ is the expected content of the b -th bin. The expected yields were evaluated by means of the Monte Carlo templates;

- n_b is the observed content of the b -th bin;
- θ is a set of nuisance parameters (NPs) which encoded the effects of the systematic uncertainties in the signal and expected background events.

The nominal fit result in terms of μ and σ_μ was obtained by maximising the likelihood function with respect to all parameters (μ and θ). The parameters which maximise the likelihood are denoted by $\hat{\mu}$ and $\hat{\theta}$. The observed signal strength μ^{obs} was the best-fit value $\hat{\mu}$, which is related to the predicted SM cross-section $\sigma_{\text{EW } VVjj}$ (multiplied by the branching ratio of the semi-leptonic decay channels), through the relation

$$\sigma_{\text{EW } VVjj}^{\text{obs}} = \mu_{\text{EW } VVjj}^{\text{obs}} \sigma_{\text{EW } VVjj} \quad (5.2)$$

so that the measurement of the signal strength could be translated to the cross-section (times the branching ratio of the decay channels observed).

5.1.2 Fit inputs

The analysis discriminating distributions were arranged as the product of channels, regions and regimes. Here, channel refers to the analysis with different number of leptons in the final states (0-lepton, 1-lepton and 2-lepton); regime refers to the merged and resolved jet analysis objects used in the selection; region refers to signal (SR) and control regions (CR). In the merged regime, the signal and control regions are divided into high-purity and low-purity subregions based on the definitions of the “tight” (50% efficiency) and “loose” (80% efficiency) working points of the boosted boson tagger. In total, 21 regions were fitted simultaneously:

- 0-lepton channel: 3 signal regions (merged high purity, merged low purity and resolved), 3 control regions (merged high purity, merged low purity and resolved regimes for Vjj);
- 1-lepton channel: 3 signal regions (merged high purity, merged low purity and resolved), 6 control regions (merged high purity, merged low purity and resolved regimes for W +jets e and top processes);
- 2-lepton channel: 3 signal regions (merged high purity, merged low purity and resolved), 3 control regions (merged high purity, merged low purity and resolved regimes for Z +jets);

The binning of the discriminant for each region was determined by requiring that statistical uncertainty in each bin be less than 5%. For each one of the signal regions, the input to the likelihood was the BDT output, while the m_{jj}^{tag} distributions were exploited as input to the likelihood function in the control regions (Table 5.1).

	0-lepton			1-lepton			2-lepton		
	Merged HP	Merged LP	Resolved	Merged HP	Merged LP	Resolved	Merged HP	Merged LP	Resolved
SR	BDT	BDT	BDT	BDT	BDT	BDT	BDT	BDT	BDT
CR Vjj	m_{jj}^{tag}	m_{jj}^{tag}	m_{jj}^{tag}	-	-	-	-	-	-
WCR	-	-	-	m_{jj}^{tag}	m_{jj}^{tag}	m_{jj}^{tag}	-	-	-
TopCR	-	-	-	one bin	one bin	one bin	-	-	-
ZCR	-	-	-	-	-	-	m_{jj}^{tag}	m_{jj}^{tag}	m_{jj}^{tag}

Table 5.1: Variables employed as inputs to the likelihood function for each channel, regime and region. “One bin” implies that a single bin without any shape information was used in the corresponding fit region and “-” means that the region was not present in the fit.

The following processes were considered in the fit:

- VBS signal: EWK $WW/WZ/ZZ$ processes;
- W +jets;
- Z +jets;
- $t\bar{t}$;
- single top: s -, t -, Wt -channel;
- diboson: QCD production of $WW/WZ/ZZ$ processes.

5.1.3 Statistical test

In order to interpret the outcome of a search, one usually assesses the level of agreement of the observed data with a given hypothesis by computing the p -value, i.e. the probability (under the assumption being made) of finding data of equal or greater incompatibility with the prediction: the latter can be regarded as excluded if the observed p -value is below a specific threshold α , namely the significance level of the *statistical hypothesis test*. The p -value is usually converted into an equivalent significance Z , defined such that the probability for a Gaussian-distributed variable of finding a value that lies more than Z standard deviations away from the mean value equals the p -value.

The tool for quantifying the compatibility between observed data and the so-called *null hypothesis* H_0 is a test statistic, i.e. a random variable built on the observed data. A commonly employed test statistic, which was also chosen in the previous run search, is based on the logarithm of the profile likelihood ratio λ_μ :

$$q_\mu = -2 \ln \lambda_\mu = -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (5.3)$$

where $\hat{\mu}$ and $\hat{\theta}$ are the observed values of μ and θ , i.e. the parameters values that maximise the likelihood function 5.1, while $\hat{\theta}_\mu$ are the values of the nuisance parameters that maximize the likelihood function for a given value of μ , i.e. they are the conditional maximum-likelihood estimators of θ and thus are functions of μ . The constraint $0 \leq \hat{\mu} \leq \mu$ is required to be fulfilled.

By definition, the profile likelihood ratio λ_μ is such that $0 \leq \lambda \leq 1$, where $\lambda = 1$ if the hypothesised μ equals the observed $\hat{\mu}$ (maximum compatibility between hypothesis and observed data). Consequently, $q_\mu = 0$ is indicative of the maximum agreement between the hypothesis and the observed data, while high (positive) values of q_μ suggest that the hypothesis may not fit the data. Quantitatively speaking, if $\text{pdf}(q_\mu | \mu)$ is the probability distribution of q_μ given the hypothesised μ , and

$$p\text{-value} = \int_{q_\mu^{\text{obs}}}^{\infty} \text{pdf}(q_\mu | \mu) dq_\mu < \alpha \quad (5.4)$$

where α is arbitrarily set *a priori*, the null hypothesis H_0 is excluded with a significance of Z standard deviations corresponding to the observed p -value (Figure 5.1).

In the previous run search, the fitted value of the signal strength turned out to be

$$\mu_{\text{EWV}Vjj}^{\text{obs}} = 1.05_{-0.40}^{+0.42} = 1.05 \pm 0.20(\text{stat.})_{-0.34}^{+0.37}(\text{syst.}) \quad (5.5)$$

and the null hypothesis H_0 to be tested was the background-only model ($\mu = 0$ hypothesis, i.e. absence of EW $VVjj$ signal). The latter was excluded in data with a significance of 2.7 standard deviations [31].

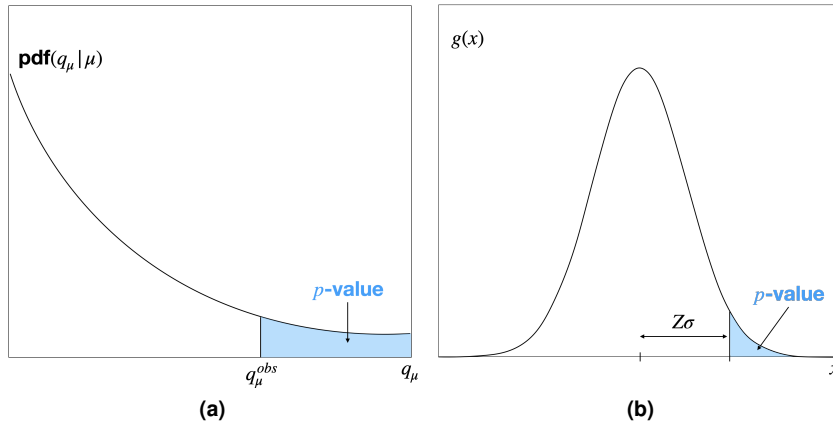


Figure 5.1: (a): Illustration of the p -value obtained from the observed value of a test statistic q_μ . (b): Relation between the p -value and the alternative Z significance.

5.2 Current approach

5.2.1 Expected significance

When planning an experiment, it is useful to quantify the sensitivity to a potential discovery. Rather than being interested in the significance obtained from a single dataset, one would want to assess the expected (more precisely, median) significance with which different values of μ (different hypotheses) could be rejected [49].

As already pointed out, in order to exclude the background-only hypothesis ($\mu = 0$) on the basis of the experiment result $\hat{\mu}$, one needs $\text{pdf}(q_\mu|\mu)$ for $\mu = 0$, i.e. $\text{pdf}(q_0|0)$, assuming data being distributed according to $\mu = 0$. Since we are now interested in the experiment sensitivity to some signal in the data (before the experiment is actually carried out), let us assume data being distributed according to a signal model described by μ' and suppose we want to test the null hypothesis described by μ (not necessarily the background-only one): we would need $\text{pdf}(q_\mu|\mu')$, where the subscript of q refers to the hypothesis being tested, while the second argument of pdf refers to the signal strength value assumed in data. The sensitivity of the experiment can be estimated by quoting the median significance, i.e. $\text{med}[q_\mu, \mu']$: the expected significance refers to the p -value corresponding to the median q_μ (rather than an observed q_μ^{oss}), assuming the alternative value μ' (Figure 5.2).

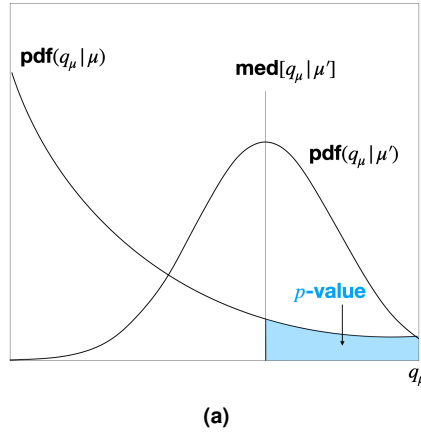


Figure 5.2: p -value corresponding to a median q_μ assuming a strength parameter μ' .

Following the procedure illustrated in [49], the median significance can be estimated by means of an Asimov dataset, i.e. a dataset such that when one uses it to evaluate the estimators for all parameters, one obtains the true parameter values. In other words, $\hat{\mu} = \mu'$, $\hat{\theta} = \theta$ and

$$q_{\mu,A} = -2 \ln \lambda_{\mu,A} = -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\mu', \theta)} \approx \frac{(\mu - \mu')^2}{\sigma^2} \quad (5.6)$$

where σ is the standard deviation of the Gaussian distribution of mean μ' which we assume $\hat{\mu}$ to follow.

As presented in the previous chapter, the current search for the electroweak $VVjj$ production will benefit from the employment of a Recurrent Neural Network. In order to estimate the sensitivity to the electroweak $VVjj$ production signal strength, the output of the RNN (trained in the 2-lepton channel signal regions, as explained in section 4.4.3) is taken into account to build the profile likelihood in reference to a multivariate approach; for the purpose of comparison, a cut-based approach is tested as well, with m_{VV} (the diboson system invariant mass) employed as final discriminant. The statistical tool used is the same as the one employed in the previous run search, namely WSMaker. Only the 1-lepton and 2-lepton channels are considered, systematic uncertainties are not taken into account and the discriminant variables distributions binning is not yet optimised. Figure 5.3 shows the profile likelihood ratio 5.6 built on an Asimov dataset, in which the assumed EW $VVjj$ signal strength is $\mu' = 1$. With reference to the plots in the figure, tighter parabolas are indicative of higher sensitivity, since even values of μ that are "close" to $\mu' = 1$ result in a high value of the profile log-likelihood ratio; in other words, an observed $\mu' = 1$ would be able to reject a

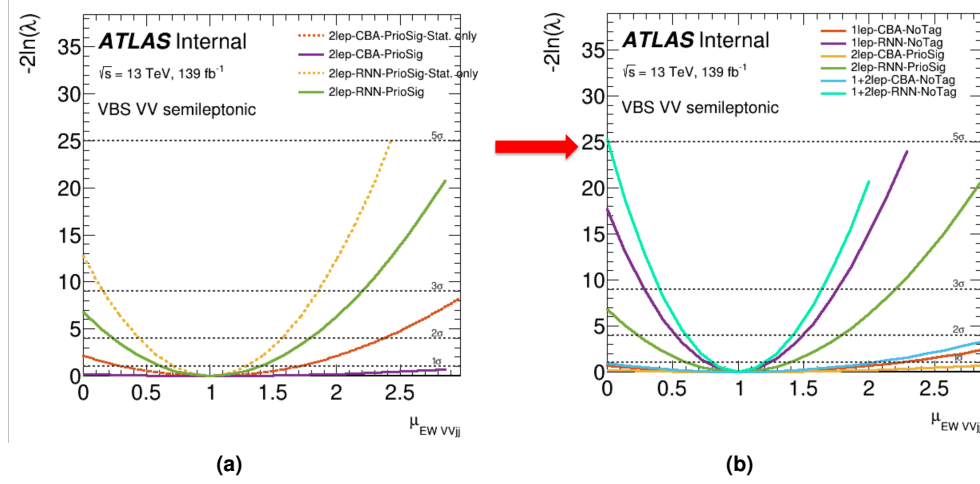


Figure 5.3: (a): Expected significance of EW $VVjj$ signal strength estimated with an Asimov dataset from the 2-lepton channel only, in the cut based approach, with (dashed red line) and without (violet line) statistical uncertainties, and in the RNN approach, with (orange dashed line) and without (green line) statistical uncertainties. (b): Expected significance of EW $VVjj$ signal strength estimated with an Asimov dataset, in the cut based approach (yellow, red and azure lines for the 1-lepton channel, 2-lepton channel and 1-lepton and 2-lepton channels together, respectively) and in the RNN approach (violet, green and turquoise lines for the 1-lepton channel, 2-lepton channel and 1-lepton and 2-lepton channels together, respectively). The red arrow points to the RNN case for 1-lepton and 2-lepton channels together (turquoise line): this setup would allow for the exclusion of the background-only hypothesis with a 5σ significance.

wide range of hypothesis (even those such that $\mu \sim 1$) with high significance. As it is made clear from the plots, the curves corresponding to the RNN approach provide a much higher sensitivity than the cut based approach; in particular, combining the 1-lepton and 2-lepton channels, the background-only model ($\mu = 0$) could be excluded with an expected significance of 5σ . This remarkable result provides the RNN approach pursued in this Thesis work with further motivation.

5.2.2 aQCGs search strategy

An excess of events with respect to the SM VBS cross-section prediction could indicate the presence of anomalous quartic gauge couplings (aQGCs). In contrast to the previous run search, the current data may provide the search for anomalous couplings with an observation of such an enhancement of the SM yields; otherwise, more stringent limits could be set on aQGCs with respect to the already observed ones [18, 23, 24, 30, 33, 37–39, 42, 43].

An excess of events could be observed in data by excluding the SM hypothesis (comprehensive of the SM electroweak $VVjj$ production). In other words, the profile log-likelihood ratio 5.3 could be exploited to exclude the $\mu = 1$ null hypothesis on the basis of the observed $\hat{\mu}$:

$$q_1 = -2 \ln \frac{\mathcal{L}(1, \hat{\theta}_1)}{\mathcal{L}(\hat{\mu}, \hat{\theta})}. \quad (5.7)$$

In addition, one could make use of the EFT prediction according to which dimension-8 operators of the EFT Lagrangian modify the SM quartic gauge couplings, leading to an enhancement of the SM VBS cross-section at high values of m_{VV} (see section 1.3.5). Instead of defining a signal strength parameter μ , the profile likelihood function can be written in terms of EFT parameters in order to perform a fit to the observed binned m_{VV} distribution. A test statistic similar to 5.3 could be defined to extract exclusion intervals derived with the CL_s method, described in [70], for f_8/Λ^4 EFT operators coefficients.

Conclusions

In this Thesis my contribution to the search for electroweak diboson production in semileptonic final states is presented. This work is part of the research performed by the ATLAS Collaboration in this field. The data exploited in the search are from the Large Hadron Collider proton-proton collisions occurred at centre-of-mass energy of 13 TeV during the full Run 2 period (from 2015 to 2018). The data were recorded with the ATLAS detector and correspond to an integrated luminosity of 139 fb^{-1} .

At the LHC, electroweak diboson production is characterised by the presence of an additional dijet system with a large invariant mass and a large pseudo-rapidity separation. In the consequent $VVjj$ final state, one vector boson V ($V = Z/W$) is searched as hadronically decaying (to improve the total branching ratio), while the leptonic decay is selected for the other boson (to exploit a clean experimental signature). The hadronically decaying boson can be identified as either two separate small-radius jets (resolved regime) or one large-radius jet (merged regime), according to the energy of the boson: the merged selection is optimised to recognize massive bosons that are highly boosted.

In this search, the diboson process of interest is vector boson scattering (VBS), which is involved with auto-interactions of the longitudinal components of the Standard Model vector bosons. However, both VBS and non-VBS electroweak processes contribute to the total electroweak diboson production cross-section, and these contribution cannot be separated in a gauge-invariant way. In consequence, the analysis goal is to measure the electroweak $VVjj$ production cross-section, exploiting an unprecedented sensitivity which could make it possible to observe or, otherwise, set more stringent limits on anomalous quartic gauge couplings. In the current search, VBS topology is identified by means of a "cut-based" approach: a pair of jets with a large pseudo-rapidity separation is searched, and its invariant mass is required to be greater than 400 GeV. Since they are employed to "tag" VBS processes, these jets are referred to as "tagging" jets.

My contribution focuses on the optimisation of the event selection as regards the 2-lepton channel, where two charged leptons (electrons or muons) are required in the final state ($ZV \rightarrow llqq$). In particular, an alternative to the cut-based approach, involving the employment of a Recurrent Neural Network, is proposed in order to retrieve signal events which are discarded due to the tagging jets invariant mass cut. Initially, the RNN was trained on Monte Carlo simulated events such that a leptonically decaying Z boson is found, exploiting information about the first six small-radius jets (p_T -ordered), with a view to identify VBS topology without searching for tagging jets at all and, in doing so, recover a significant amount of signal efficiency. Successively, the RNN training was performed in the signal regions, i.e. using Monte Carlo events such that, in addition to the leptonically decaying Z candidate, one or two (depending on the considered regime) candidate signal jets from the hadronic decay of Z or W are found. In order to carry out the training, information about both the small-radius jets and the large-radius jets were taken into account. The latter strategy results in a remarkable enhancement of the expected significance, while keeping a high expected signal efficiency with respect to the cut-based approach. Indeed, evaluating the sensitivity of the analysis to the electroweak $VVjj$ production signal, the expected significance with which the Standard Model background-only hypothesis could be excluded is about 5σ when the RNN is employed to discriminate signal events against background ones.

Bibliography

- [1] URL: <http://tmva.sourceforge.net>.
- [2] Gerardus 't Hooft. “Renormalization of massless Yang-Mills fields”. In: *Nuclear physics: B* 33.1 (1971), pp. 173–199. DOI: 10.1016/0550-3213(71)90395-6.
- [3] E Abat et al. “Combined performance studies for electrons at the 2004 ATLAS combined test-beam”. In: *Journal of Instrumentation* 5.11 (2010), P11006. DOI: 10.1088/1748-0221/5/11/P11006.
- [4] Sea Agostinelli et al. “GEANT4—a simulation toolkit”. In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [5] M Aharrouche et al. “Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 614.3 (2010), pp. 400–432. DOI: 10.1016/j.nima.2009.12.055.
- [6] F. Chollet et al. *Keras*. GitHub repository. 2015. URL: <https://github.com/fchollet/keras>.
- [7] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *Journal of High Energy Physics* 2010.6 (2010), p. 43. DOI: 10.1007/JHEP06(2010)043.
- [8] Johan 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: *Journal of High Energy Physics* 2014.7 (2014), p. 79. DOI: 10.1007/JHEP07(2014)079.
- [9] G Arnison et al. “Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}=540$ GeV”. In: *Physics letters B* 122.1 (1983), pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2.
- [10] G Arnison et al. “Experimental observation of lepton pairs of invariant mass around 95 GeV/ c^2 at the CERN SPS collider”. In: *Physics Letters B* 126.5 (1983), pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0.
- [11] *ATLAS Pythia 8 tunes to 7 TeV datas*. Tech. rep. ATL-PHYS-PUB-2014-021. Geneva: CERN, Nov. 2014. URL: <https://cds.cern.ch/record/1966419>.
- [12] V Barger et al. “Strong W^+W^+ scattering signals at pp supercolliders”. In: *Physical Review D* 42.9 (1990), p. 3052. DOI: 10.1103/PhysRevD.42.3052.
- [13] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “FastJet user manual”. In: *The European Physical Journal C* 72.3 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2.
- [14] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti- k_t jet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063.
- [15] Matteo Cacciari and Gavin P. Salam. “Pileup subtraction using jet areas”. In: *Physics Letters B* 659.1-2 (Jan. 2008), pp. 119–126. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2007.09.077.
- [16] ATLAS Collaboration. “Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data”. In: *The European Physical Journal C* 72.3 (2012), p. 1909. DOI: 10.1140/epjc/s10052-012-1909-1.

- [17] ATLAS Collaboration. “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 (Aug. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6.
- [18] ATLAS Collaboration. “Evidence for Electroweak Production of $W^\pm W^\pm jj$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector”. In: *Phys. Rev. Lett.* 113 (14 Oct. 2014), p. 141803. DOI: 10.1103/PhysRevLett.113.141803.
- [19] ATLAS Collaboration. “Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV”. In: *The European Physical Journal C* 73.3 (2013), p. 2304. DOI: 10.1140/epjc/s10052-013-2304-2.
- [20] ATLAS Collaboration. *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. July 2020. arXiv: 2007.02645v1 [hep-ex].
- [21] ATLAS Collaboration. “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), pp. 1–47. DOI: 10.1007/JHEP05(2012)128.
- [22] ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *The European Physical Journal C* 77.7 (July 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5031-2.
- [23] ATLAS Collaboration. “Measurements of $W^\pm Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings”. In: *Phys. Rev. D* 93 (9 May 2016), p. 092004. DOI: 10.1103/PhysRevD.93.092004.
- [24] ATLAS Collaboration. “Measurements of electroweak Wjj production and constraints on anomalous gauge couplings with the ATLAS detector”. In: *The European Physical Journal C* 77.7 (2017), p. 474. DOI: 10.1140/epjc/s10052-017-5007-2.
- [25] ATLAS Collaboration. *Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV*. 2020. arXiv: 2004.03969 [hep-ex].
- [26] ATLAS Collaboration. “Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment”. In: *The European Physical Journal C* 76.1 (2016), pp. 1–51. DOI: 10.1140/epjc/s10052-015-3769-y.
- [27] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y.
- [28] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.
- [29] ATLAS Collaboration. “Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector”. In: *The European Physical Journal C* 76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z.
- [30] ATLAS Collaboration. “Search for anomalous electroweak production of WW/WZ in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Phys. Rev. D* 95 (3 Feb. 2017), p. 032001. DOI: 10.1103/PhysRevD.95.032001.
- [31] ATLAS Collaboration. “Search for 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: *Physical Review D* 100.3 (2019), p. 032007. DOI: 10.1103/PhysRevD.100.032007.
- [32] ATLAS Collaboration. “Search for heavy diboson resonances in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: (2020). arXiv: 2004.14636.
- [33] ATLAS Collaboration. “Studies of $Z\gamma$ production in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2017.7 (2017), p. 107. DOI: 10.1007/JHEP07(2017)107.

- [34] ATLAS Collaboration. “Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector”. In: *The European Physical Journal C* 75.10 (2015), p. 476. DOI: 10.1140/epjc/s10052-015-3685-1.
- [35] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *The European Physical Journal C* 70.3 (Sept. 2010), pp. 823–874. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1429-9.
- [36] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5.
- [37] CMS Collaboration. *Evidence for exclusive $\gamma\gamma \rightarrow W^+W^-$ production and constraints on anomalous quartic gauge couplings in pp collisions at $\sqrt{s} = 7$ and 8 TeV*. Tech. rep. Springer/SISSA, 2016. arXiv: 1604.04464v1 [hep-ex].
- [38] CMS Collaboration. “Measurement of electroweak-induced production of $W\gamma$ with two jets in pp collisions at $\sqrt{s} = 8$ TeV and constraints on anomalous quartic gauge couplings”. In: *Journal of High Energy Physics* 2017.6 (2017), p. 106. DOI: 10.1007/JHEP06(2017)106.
- [39] CMS Collaboration. “Measurement of the cross section for electroweak production of $Z\gamma$ in association with two jets and constraints on anomalous quartic gauge couplings in proton–proton collisions at $\sqrt{s} = 8$ TeV”. In: *Physics Letters B* 770 (2017), pp. 380–402. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2017.04.071.
- [40] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021.
- [41] CMS Collaboration. “Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV”. In: *The European Physical Journal C* 75.5 (May 2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3351-7.
- [42] CMS Collaboration. “Search for anomalous electroweak production of vector boson pairs in association with two jets in proton-proton collisions at 13 TeV”. In: *Physics Letters B* 798 (2019), p. 134985. DOI: 10.1016/j.physletb.2019.134985.
- [43] CMS Collaboration. “Study of Vector Boson Scattering and Search for New Physics in Events with Two Same-Sign Leptons and Two Jets”. In: *Phys. Rev. Lett.* 114 (5 Feb. 2015), p. 051801. DOI: 10.1103/PhysRevLett.114.051801.
- [44] NA48 Collaboration. “A New measurement of direct CP violation in two pion decays of the neutral kaon”. In: *Physics Letters B* 465.1-4 (1999), pp. 335–348. DOI: 10.1016/S0370-2693(99)01030-8.
- [45] UA2 Collaboration. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp Collider”. In: *Physics Letters B* 129.1-2 (1983), pp. 130–140. DOI: 10.1016/0370-2693(83)90744-X.
- [46] UA2 Collaboration. “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $\bar{p}p$ collider”. In: *Phys. Lett. B* 122.CERN-EP-83-25 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.
- [47] “Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV”. In: *Phys. Rev. D* 92 (1 July 2015), p. 012004. DOI: 10.1103/PhysRevD.92.012004.
- [48] Thijs G Cornelissen et al. “The global χ^2 track fitter in ATLAS”. In: *Journal of Physics: Conference Series*. Vol. 119. 3. IOP Publishing, 2008, p. 032013. DOI: 10.1088/1742-6596/119/3/032013.
- [49] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (2011), pp. 1–19. DOI: 10.1140/epjc/s10052-011-1554-0. Corrected in: “Erratum to: Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 73.7 (2013), p. 2501. DOI: 10.1140/epjc/s10052-013-2501-z.

- [50] Céline Degrande et al. “Effective field theory: A modern approach to anomalous couplings”. In: *Annals of Physics* 335 (Aug. 2013), pp. 21–32. ISSN: 0003-4916. DOI: 10.1016/j.aop.2013.04.016.
- [51] O. J. P. Éboli and M. C. Gonzalez–Garcia. “Classifying the bosonic quartic couplings”. In: *Physical Review D* 93.9 (May 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.093013.
- [52] François Englert and Robert Brout. “Broken symmetry and the mass of gauge vector mesons”. In: *Physical Review Letters* 13.9 (1964), p. 321. DOI: 10.1103/PhysRevLett.13.321.
- [53] *Expected electron performance in the ATLAS experiment*. Tech. rep. ATL-PHYS-PUB-2011-006. Geneva: CERN, Apr. 2011. URL: <http://cds.cern.ch/record/1345327>.
- [54] Rudolf Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2-3 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.
- [55] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear physics* 22.4 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.
- [56] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *Journal of High Energy Physics* 2009.02 (Feb. 2009), pp. 007–007. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2009/02/007.
- [57] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. “Broken Symmetries”. In: *Phys. Rev.* 127 (3 Aug. 1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965.
- [58] O. W. Greenberg. “Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons”. In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 598–602. DOI: 10.1103/PhysRevLett.13.598.
- [59] M. Y. Han and Y. Nambu. “Three-Triplet Model with Double SU(3) Symmetry”. In: *Phys. Rev.* 139 (4B Aug. 1965), B1006–B1010. DOI: 10.1103/PhysRev.139.B1006.
- [60] FJ Hasert et al. “Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment”. In: *Nuclear Physics B* 73.1 (1974), pp. 1–22. DOI: 10.1016/0370-2693(73)90499-1.
- [61] Peter W Higgs. “Broken symmetries and the masses of gauge bosons”. In: *Physical Review Letters* 13.16 (1964), p. 508. DOI: 10.1103/PhysRevLett.13.508.
- [62] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012. URL: <http://cds.cern.ch/record/1449796>.
- [63] *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. Tech. rep. ATLAS-CONF-2016-035. Geneva: CERN, July 2016. URL: <https://cds.cern.ch/record/2200211>.
- [64] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jet trimming”. In: *Journal of High Energy Physics* 2010.2 (2010), p. 84. DOI: 10.1007/JHEP02(2010)084.
- [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, Apr. 2008. URL: <http://cds.cern.ch/record/1099735>.
- [66] Andrew J. Larkoski, Ian Moulton, and Duff Neill. “Power counting to better jet observables”. In: *Journal of High Energy Physics* 2014.12 (Dec. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep12(2014)009.
- [67] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. “Energy correlation functions for jet substructure”. In: *Journal of High Energy Physics* 2013.6 (June 2013). ISSN: 1029-8479. DOI: 10.1007/jhep06(2013)108.
- [68] 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 Apr. 1977), pp. 883–885. DOI: 10.1103/PhysRevLett.38.883.
- [69] Benjamin W. Lee, C. Quigg, and H. B. Thacker. “Weak interactions at very high energies: The role of the Higgs-boson mass”. In: *Phys. Rev. D* 16 (5 Sept. 1977), pp. 1519–1531. DOI: 10.1103/PhysRevD.16.1519.

- [70] Alexander L Read. “Presentation of search results: the CL_s technique”. In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313.
- [71] A Salam. “N. Svartholm, ed. Elementary Particle Physics: Relativistic Groups and Analyticity”. In: *Eighth Nobel Symposium. Stockholm: Almqvist and Wiksell*. 1968, p. 367.
- [72] Andreas Salzburger. “Track Simulation and Reconstruction in the ATLAS experiment”. 2008. URL: <http://cds.cern.ch/record/2224514>.
- [73] Torbjörn Sjöstrand. *PYTHIA 8 Status Report*. 2008. arXiv: 0809.0303 [hep-ph].
- [74] The Theano Development Team et al. *Theano: A Python framework for fast computation of mathematical expressions*. 2016. arXiv: 1605.02688 [cs.SC].
- [75] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.
- [76] C. N. Yang and R. L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. In: *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191.