Università degli Studi di Napoli “Federico II”

Dottorato di Ricerca in Fisica

Ciclo XXXV

Coordinatore: Prof. Vincenzo Canale

Improving the search of new resonances Y decaying into a Higgs boson and a generic new particle X in hadronic final states with Machine Learning at the ATLAS detector

Settore Scientifico Disciplinare: Fis/04

Tutori:
Prof. Francesco Conventi
Prof. Leonardo Merola
Prof. Elvira Rossi

Dottorando:
Silvia Auricchio

Improving the search of new resonances Y decaying into a Higgs boson and a generic new particle X in hadronic final states with Machine Learning at the ATLAS detector

University of Naples "Federico II" - Department of Physics

Abstract

High Energy Physics is now accessing a new era. The research for Physics Beyond Standard Model requires more and more data to prove the existence of eventual rare phenomena. This involves a big challenge on the processing of the considerable amount and complexity of data collected and Machine learning (ML) provides a valid solution. The work presented here is about the search for a heavy resonance Y decaying into a Standard Model Higgs boson H and a new particle X in a fully hadronic final state, with several ML applications which helped to improve several aspects of the analysis. The results presented are based on the 139 fb^{-1} dataset of proton-proton collisions at $\sqrt{s}$ = 13 TeV, collected by the ATLAS detector from 2015 to 2018, during LHC Run-2. A novel anomaly detection signal region is implemented based on a jet-level score for signal model-independent tagging of the boosted X, representing the first application of fully unsupervised ML to an ATLAS analysis. Two additional signal regions are implemented to target a benchmark X decay to two quarks, covering topologies where the X is reconstructed as either a single large-radius jet or two small-radius jets. The Higgs is assumed to decay to $b\bar{b}$ and its boosted topology is recognized among Quantum Chromodynamics and Top jets using a new ML-based tagger. The background estimation is totally data-driven and performed with the help of a Deep Neural Network. No significant excess of data is observed over the expected background, and the results are interpreted in upper limits at 95% confidence level on the production cross section $\sigma(pp \to Y \to XH \to q\bar{q}b\bar{b})$ for signals with m_Y between 1.5 and 6 TeV and m_X between 65 and 3000 GeV.

Contents

Introduction

The success of the Standard Model (SM) in describing the wide range of precise experimental measurements is a remarkable achievement. It is constructed from a number of beautiful and profound theoretical ideas put together in order to reproduce the experimental data. However, many mysteries remain, such as the nature of dark matter and the origin of matter-antimatter asymmetry in the universe, which confirm the need for new physics beyond the Standard Model (BSM).
Several theories have been formulated, which describe the SM predictions at higher energy and many of them predict the existence of heavy resonances.

The Large Hadron Collider (LHC), as the world's highest energy proton-proton (pp) collider, is a unique facility for the search of these hypothetical new particles. This thesis documents the work realized with 139 fb^{-1}of *pp* collisions data collected by the ATLAS detector during LHC Run-2, looking for final states with two large-R jets, coming from the decay of a new TeV-scale narrow-width heavy resonance Y into a SM Higgs boson and another particle X, with the mass varying from $\mathcal{O}(10)$ GeV to $\mathcal{O}(1)$ TeV.

A search for the $Y \to XH$ process was previously performed by ATLAS using 36.1 fb^{-1} under the assumption of $X \to q\bar{q}$, with no significant excess found covering Y masses from 1 to 4 TeV and X masses from 50 to 1000 GeV [32]. In addition to the increased luminosity of the dataset, the current round of the analysis includes several key improvements with respect to this last iteration, such as a neural net-based tagger optimized for the boosted $H \to b\bar{b}$ topology, anomaly detection for enhanced signal model independence, a deep neural network-based method for the background estimation and the usage of two orthogonal regions to capture both boosted and resolved reconstruction of the nominal X decay to two quarks.

This thesis consists of four chapters. In Chapter 1 an overview of the Standard Model and the Higgs mechanism is presented; in Chapter 2 the ATLAS detector with all its sub-constituents is described, with additional details about the ATLAS trigger system and the LHC long term schedule; Chapter 3 is about the reconstruction of the main physics objects; finally, in Chapter 4 the $Y \to XH$ analysis is described in detail and the obtained results are discussed. In addition, two Appendices with insights on the ATLAS New Small Wheels and on several aspects of the analysis are present at the end of the Thesis.

CHAPTER 1

The Standard Model of the elementary particles

All of nature springs from a handful of components, the *fundamental particles*, that interact with one another via *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. In the 1970s, physicists developed a set of equations describing these particles and interactions (excluding gravity). Together, the equations formed a succinct theory known as the Standard Model (SM) of particle physics.

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, referred to as *internal* since they do not deal with the kinematic state of the particle itself. One of these numbers is the *spin* (whose value shall be from now on expressed in units of Planck's constant $\hbar$) and, according to its value, elementary particles can be split into two fundamental categories:

- 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.

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 $e \approx 1.6 \cdot 10^{-15}$ C).

Electrons are fermions with $s = 1/2$ and belong to the lepton class. Looking at scales smaller then atom's, 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).
Fermions are further classified into:

- **Leptons** grouped in three families, one for each flavor, each comprising a charged lepton (e^-, μ^-, τ^-) and a weakly interactive neutral particle, called neutrino (ν_e, ν_μ, ν_τ).
- **Quarks** existing in six different "flavours" and grouped in three families.

1.1.1 Leptons and quarks

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 $L = -1$, that is always conserved in any physical processes. 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.
Similarly to leptons, quarks can interact electromagnetically (being provided with fractional electric charges) and weakly, but the difference with the latter is that they also interact strongly. 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 1.1 (including the antiparticles as well).
Quarks are given a barion 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).

1.1.2 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

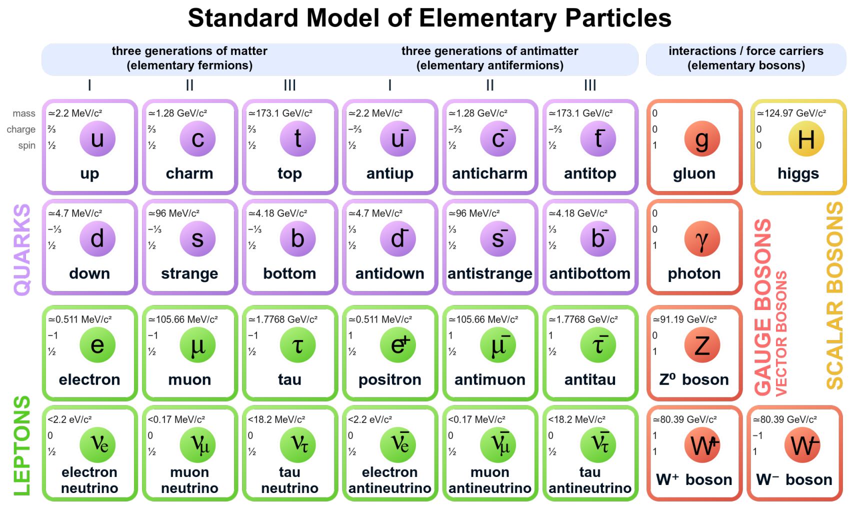


Figure 1.1: The elementary particles and antiparticles, divided into quarks, leptons and force-carrying particles, together with the scalar Higgs boson.

particle-like objects displayed a wave-like behaviour as well.

Whereas matter and contact forces were studied on a intuitive basis, the nature of interactions between distant bodies was always problematic to deal with. Only with the introduction of *field* concept it found an explanation. A classical field is a function assigning a vector at each point of space and time, and every (gravitational or charged) body ”feels” the presence of the field, which is responsible of the (gravitational or electrical) force to act on it.

At present, interactions between particles are not described by classical fields any more: relativistic *Quantum Field Theory* describes the interactions by the means of quantum fields. 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 can a particle 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 rep-

resented in Figure 1.2 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.

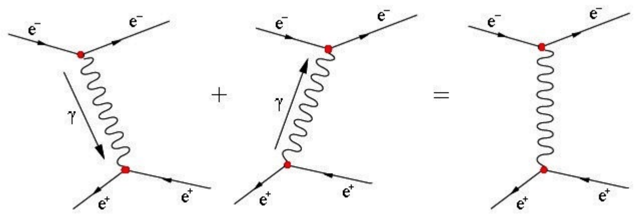


Figure 1.2: The electromagnetic interaction between two electrons by exchange of a photon.

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 1.1, 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 Standard Model

Bearing in mind the local gauge invariance principle [45] 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 1.2.2) 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 $G = \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y \otimes \mathrm{SU}(3)_c$. In particular:

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

1.2.1 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 [80]. Later on, Steven Weinberg [111] and Abdus Salam [105] integrated the Brout–Englert–Higgs mechanism in Glashow's theory, giving birth to the modern GWS model.
In $\mathrm{SU}(2)_L\otimes \mathrm{U}(1)_Y$, $\mathrm{SU}(2)_L$ 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:

$$P_L = \frac{1}{2}(1-\gamma^5)\,, \quad P_R = \frac{1}{2}(1+\gamma^5) \tag{1.1}$$

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

$$T_i = \frac{\sigma_i}{2}\,, \quad [\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \tag{1.2}$$

where the Levi-Civita symbol ε_{ijk} provides 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 SU(2)$_L$, the left-handed components of the fermion fields are organized into doublets which undergo SU(2)$_L$ transformations, while the right-handed components remain unchanged and are therefore arranged into weak isospin singlets:

$$\mathrm{SU}(2)_L: \quad \left(\begin{smallmatrix} f_1 \\ f_2 \end{smallmatrix}\right)'_L = e^{-ig\frac{\boldsymbol{\sigma}}{2}\cdot\boldsymbol{\theta}(x)} \left(\begin{smallmatrix} f_1 \\ f_2 \end{smallmatrix}\right)_L, \quad f'_R = f_R \tag{1.3}$$

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

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.4}$$

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). 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

$$I_3 + \frac{Y}{2} = Q \tag{1.5}$$

gives rise to the generator Q of the electromagnetic interaction symmetry group U(1)$_{\mathrm{em}}$. 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)$_{\mathrm{em}}$ generators must be taken into account.

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

$$\mathrm{U}(1)_Y: \quad \left(\begin{smallmatrix} f_1 \\ f_2 \end{smallmatrix}\right)'_L = e^{-ig'Y\delta(x)} \left(\begin{smallmatrix} f_1 \\ f_2 \end{smallmatrix}\right)_L, \quad f'_R = e^{-ig'Y\delta(x)} f_R \tag{1.6}$$

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, the free Lagrangian is

$$\mathcal{L} = i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + i\bar{\psi}_R \gamma^\mu \partial_\mu \psi_R + \text{h.c.} \tag{1.7}$$

where:

- ψ_L and ψ_R are a left-handed doublet and a right-handed singlet of SU(2)$_L$, respectively;

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

- the sum over all the fermions is implied;
- h.c. stands for the hermitian conjugate and takes account of the antifermions.

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

$$\mathcal{D}^L_\mu = \partial_\mu + ig\frac{\boldsymbol{\sigma}}{2}\cdot \mathbf{W}_\mu + ig'YB_\mu\,, \quad \mathcal{D}^R_\mu = \partial_\mu + ig'YB_\mu \tag{1.8}$$

where $\mathcal{D}^L_\mu$ and $\mathcal{D}^R_\mu$ act on doublets and singlets, respectively. While the gauge fields $\mathbf{W}_\mu$ satisfy the condition 1.9 with respect to $\mathrm{SU}(2)_L$, the gauge field B_μ changes like 1.10 under $\mathrm{U}(1)_Y$.

$$W'^{(k)}_\mu = W^{(k)}_\mu + \partial_\mu \epsilon_k - g f_{ijk}\epsilon_i W^{(j)}_\mu \tag{1.9}$$

$$B_\mu{}' = B_\mu + \partial_\mu \theta\,. \tag{1.10}$$

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

$$\mathcal{L} = i\bar{\psi_L}\gamma^\mu \mathcal{D}^L_\mu \psi_L + i\bar{\psi_R}\gamma^\mu \mathcal{D}^R_\mu \psi_R - \frac{1}{4}C^{\mu\nu}C_{\mu\nu} - \frac{1}{4}\mathbf{G}^{\mu\nu}\cdot\mathbf{G}_{\mu\nu} \tag{1.11}$$

where $C_{\mu\nu}$ and $\mathbf{G}_{\mu\nu}$ are tensors which satisfy the relations:

$$C_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.12}$$

$$G^{(i)}_{\mu\nu} = \partial_\mu W^{(i)}_\nu - \partial_\nu W^{(i)}_\mu - g(\mathbf{W}_\mu \times \mathbf{W}_\nu)^{(i)} \tag{1.13}$$

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.

1.2.2 The Brout–Englert–Higgs mechanism

In 1962, Jeffrey Goldstone, Salam and Weinberg showed [82] 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 [86] and, independently, Robert Brout and Francois Englert [76] 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

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \tag{1.14}$$

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

$$\mathcal{L} = (\mathcal{D}_\mu\phi)^*(\mathcal{D}^\mu\phi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - V(\phi) \tag{1.15}$$

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

$$v^2 = 2\phi^*_{\text{vac}}\phi_{\text{vac}} = -\frac{\mu^2}{\lambda} \tag{1.16}$$

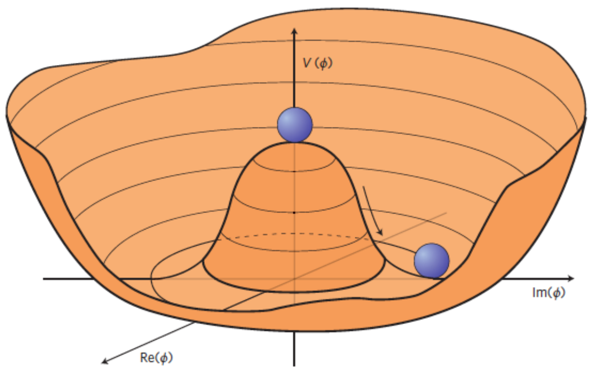


Figure 1.3: Higgs potential for a complex scalar field. Choosing any of the points at the bottom of the potential breaks spontaneously the rotational U(1) symmetry. [75]

One can therefore write ϕ_{vac} as

$$\phi_{\text{vac}} = \frac{v}{\sqrt{2}}e^{i\alpha} \tag{1.17}$$

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 1.17 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

$$\phi = \frac{v + H(x)}{\sqrt{2}}e^{i\frac{\theta(x)}{v}} \tag{1.18}$$

where $H(x)$ and $\theta(x)$ take account of the radial and phase oscillations around $\phi_{\text{vac}} = v/\sqrt{2}$, respectively (Figure 1.3). Substituting 1.18 in 1.15, 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

$$\phi' = e^{-iq\beta(x)}\phi = \frac{(v + H(x))}{\sqrt{2}} \tag{1.19}$$

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)

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.20}$$

by adding the term

$$\mathcal{L}_H = (\mathcal{D}^L_\mu \phi)^\dagger (\mathcal{D}^{L\mu}\phi) - V(\phi) \tag{1.21}$$

to the Lagrangian 1.11. Here, $\mathcal{D}^L_\mu$ is the covariant derivative 1.8, while the potential $V(\phi)$ is

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{4}(\phi^\dagger\phi)^2 \tag{1.22}$$

Again, picking one quadruplet $\phi_1, \cdots, \phi_4$ 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):

$$\begin{aligned}\mathcal{L}_{G\phi} =& \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 + \\ & - \frac{1}{4}(\partial_\mu W^{(1)}_\nu - \partial_\nu W^{(1)}_\mu)(\partial^\mu W^{(1)^\nu} - \partial^\nu W^{(1)^\mu}) + \frac{1}{8} g^2 v^2 W^{(1)}_\mu W^{(1)^\mu} + \\ & - \frac{1}{4}(\partial_\mu W^{(2)}_\nu - \partial_\nu W^{(2)}_\mu)(\partial^\mu W^{(2)^\nu} - \partial^\nu W^{(2)^\mu}) + \frac{1}{8} g^2 v^2 W^{(2)}_\mu W^{(2)^\mu} + \\ & - \frac{1}{4}(\partial_\mu W^{(3)}_\nu - \partial_\nu W^{(3)}_\mu)(\partial^\mu W^{(3)^\nu} - \partial^\nu W^{(3)^\mu}) - \frac{1}{4} C^{\mu\nu} C_{\mu\nu} + \\ & + \frac{1}{8} v^2 (g W^{(3)}_\mu - g' B_\mu)(g W^{(3)^\mu} - g' B^\mu)\end{aligned} \tag{1.23}$$

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 $W^{(1)}_\mu$ and $W^{(2)}_\mu$ ($M_W = gv/2$); the fourth line takes account of $W^{(3)}_\mu$ and B_μ kinetic terms, while the last line is a massive term of a combination of $W^{(3)}_\mu$ and B_μ. Given the Weinberg angle θ_W such that

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}\,, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \tag{1.24}$$

one can define the gauge fields Z_μ and A_μ

$$Z_\mu = W^{(3)}_\mu \cos\theta_W - B_\mu \sin\theta_W \tag{1.25}$$
$$A_\mu = W^{(3)}_\mu \sin\theta_W + B_\mu \cos\theta_W \tag{1.26}$$

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, $W^{(1)}_\mu$ and $W^{(2)}_\mu$ turn out to be mixed into two fields $W^{(+)}_\mu$ and $W^{(-)}_\mu$ (with the same mass as before) such that

$$W^{(\pm)}_\mu = \frac{W^{(1)}_\mu \mp iW^{(2)}_\mu}{\sqrt{2}} \tag{1.27}$$

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

$$q = g\sin\theta_W\,. \tag{1.28}$$

Equation 1.28 encodes the essence of the electroweak unification.

In summary, 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 [85]; ten years later, the UA1 [26, 27] and the UA2 [68, 67] collaborations provided the observation of the W and Z bosons at the Super Proton Synchroton.

In 2012, the ATLAS and CMS Collaboration announced the observation of a particle with a $125\,\mathrm{GeV}$ mass compatible with the Standard Model Higgs boson [50, 64]. This discovery earned Higgs and Englert the Nobel Prize in Physics in 2013.

1.2.3 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 [83] and Han-Nambu [84], independently. They theorized three different "values" for the color charge, so that the field representing a quark of flavour f would be

$$\psi^f(x) = \begin{pmatrix} \psi^f_R(x) \\ \psi^f_G(x) \\ \psi^f_B(x) \end{pmatrix} \tag{1.29}$$

with $\psi^f_i(x)$ 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:

$$\psi'^f_i(x) = e^{ig_s \frac{\boldsymbol{\lambda}}{2} \cdot \boldsymbol{\theta}} \psi^f_i(x) \tag{1.30}$$

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 $A^{(i)}_\mu(x)$, which are directly associated with the strong interaction massless carriers, the gluons. The QCD Lagrangian turns out to be

$$\mathcal{L}_{\text{QCD}} = i\bar{\psi}^f_i(x)\gamma^\mu \mathcal{D}_\mu \psi^f_i(x) - \frac{1}{4}\mathbf{G}^{\mu\nu} \cdot \mathbf{G}_{\mu\nu} \tag{1.31}$$

where

$$\mathcal{D}_\mu = \partial_\mu + ig_s \frac{\boldsymbol{\lambda}}{2} \cdot \mathbf{A}_\mu \tag{1.32}$$

$$G^{(i)}_{\mu\nu} = \partial_\mu A^{(i)}_\nu - \partial_\nu A^{(i)}_\mu - g_s f_{ijk} A^{(j)}_\mu A^{(k)}_\nu \tag{1.33}$$

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 1.33 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 chromodynamics: 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. 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.

1.2.4 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. This term can actually be written as

$$m\psi\bar{\psi} = m(\psi_L\bar{\psi}_R + \psi_R\bar{\psi}_L) \tag{1.34}$$

Even though equation 1.34 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 1.34 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

$$\mathcal{L}_{\text{Yukawa}} = -G_l^{ij}\bar{L}_L^i\phi l_R^j - G_d^{ij}\bar{Q}_L^i\phi d_R^j - G_u^{ij}\bar{Q}_L^i\phi_C u_R^j + \text{ h.c.} \tag{1.35}$$

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 1.20 and ϕ_C is the conjugate SU(2)$_L$ doublet of ϕ, defined as

$$\phi_C = i\sigma_2\phi^* = \begin{pmatrix}\phi^{0*}\\ -\phi^-\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_3 - i\phi_4\\ -(\phi_1 - i\phi_2)\end{pmatrix} \tag{1.36}$$

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

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}}^{\text{lepton}} &= -\left(1 + \frac{H}{v}\right)\left[\bar{l}_L^i m_l^i l_R^i + \bar{l}_R^i m_l^i l_L^i + \text{h.c.}\right] = \\ &= -m_l^i \bar{l}^i l^i - \frac{g m_l^i}{2m_W}\bar{l}^i l^i H + \text{h.c.}\end{aligned} \tag{1.37}$$

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

$$\mathcal{L}_{\text{Yukawa}}^{\text{quark}} = -\left(1 + \frac{H}{v}\right)\left[\bar{u}_L^i M_u^{ij} u_R^j + \bar{d}_L^i M_d^{ij} d_R^j + \text{h.c.}\right] \tag{1.38}$$

where

$$M_d^{ij} = G_d^{ij}\frac{v}{\sqrt{2}}, \quad M_u^{ij} = G_u^{ij}\frac{v}{\sqrt{2}} \tag{1.39}$$

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

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} = -m_l^i\bar{l}^i l^i - \frac{g m_l^i}{2m_W}\bar{l}^i l^i H - m_d^\alpha \bar{d}^\alpha d^\alpha - \frac{g m_d^\alpha}{2m_W}\bar{d}^\alpha d^\alpha H +\\ -m_u^\alpha \bar{u}^\alpha u^\alpha - \frac{g m_u^\alpha}{2m_W}\bar{u}^\alpha u^\alpha H + \text{h.c.}\end{aligned} \tag{1.40}$$

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:

$$u_L^\alpha = (U_u^L)^{\alpha i} u_L^i \quad u_R^\alpha = (U_u^R)^{\alpha i} u_R^i \tag{1.41}$$

$$d_L^\alpha = (U_d^L)^{\alpha i} d_L^i \quad d_R^\alpha = (U_d^R)^{\alpha i} d_R^i \tag{1.42}$$

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:

$$V^{\alpha\beta} = \left[U_u^L U_d^{L\dagger} \right]^{\alpha\beta} \tag{1.43}$$

As it is made clear from equation 1.40, 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 Beyond Standard Model

The SM has proven over the years to be an incredibly solid theory capable to resist many decades of continuous attempts to find its flaws. Nevertheless it is now clear it is not a complete description of how Nature works and some of the reasons are summarized in this section.
Higgs Mass Hierarchy Problem
A first hint of the incompleteness of the SM is given by the hierarchy problem, which underlines the inability of the model to homogeneously include gravity. Being a scalar particle, the Higgs boson mass is subject to radiative corrections beyond LO from all particles coupled to it. These corrections Δm_h (diagram in Figure 1.4) are in the form:

$$\Delta m_h^2 = \frac{\lambda_f}{16\pi^2} \left(2\Lambda^2 + \mathcal{O}\left[m_f^2 \ln \frac{\Lambda}{m_f} \right] \right) \tag{1.44}$$

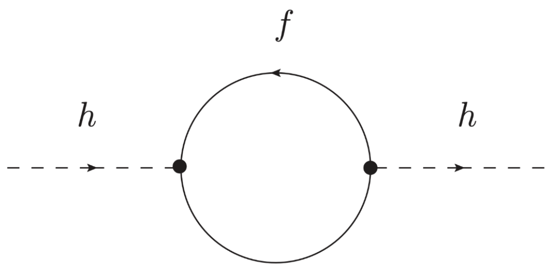


Figure 1.4: Diagram representing fermion loop contributing to Higgs mass.

with λ_f being the fermions Yukawa coupling, m_f their mass, and Λ the scale up to which SM validity holds. The mass correction is quadratically divergent in Λ. If the SM holds up to the Planck scale ($\mathcal{O}(10^{19})$, and no other phenomena come into play, to obtain the relatively low value observed for the mass of the Higgs boson, a particularly fine tuning in the cancellation of different huge terms is necessary. This fine tuning seems unnatural and unlikely, even when the anthropic principle is brought up. The problem would be solved by the discovery of new physics at a scale, assuming $\Lambda \sim \mathcal{O}(1)$ (still undetected), or proving an extended theory capable to justify it (like the Supersymmetry theory [73]).

Yukawa Couplings Hierarchy

Another problem is found in the vast hierarchy of fermions Yukawa couplings. Table 1.2 reports the value of the couplings divided by the Higgs vacuum expectation value v. This structure with dimensionless parameters spanning six orders

Table 1.2: Fermions Yukawa couplings in the standard models, divided by Higgs vacuum expectation value ($\sim$ 246), grouped by family.

First Family		Second Family		Third Family	
y_u	10^{-5}	y_c	7.3×10^{-3}	y_t	~ 1
y_d	3×10^{-5}	y_s	6×10^{-4}	y_b	2.4×10^{-2}
y_e	2.9×10^{-6}	y_μ	6.1×10^{-4}	y_τ	1×10^{-2}

of magnitude requires some kind of explanation, not provided by the SM. Many theories have been proposed to justify this phenomenon, but yet to be proven right.

Neutrino Masses

An additional omission of the SM, but this time at low energies, is the non-zero mass of neutrinos, demonstrated through their flavour oscillations. This opens the question of how to obtain a mass so small in a natural way. A Yukawa coupling to achieve this mass would make the hierarchy another factor of 10^6 wider, and the same kind of coupling as the other fermions assumes the existence of a right-handed neutrino [96] which has yet to be observed.

Dark Matter

Many independent astrophysical observations have proven the existence in the universe of a large quantity of energy density which acts as ordinary matter in terms of gravitational effects, but interacts with SM particles in no other noticeable way. Indeed this component seems to be five times more abundant than ordinary matter. Since not emitting, it has been called Dark Matter (DM)[78].

The observed DM abundance in the early Universe can arise from different production mechanisms depending on the DM coupling to the SM. In the case where the DM reached a thermal equilibrium with the SM, the production can be explained by a freeze-out mechanism. It consists of the particle density freezing when the interaction rate between DM and SM particles cannot compensate anymore the Hubble expansion rate. Considering a DM particle χ and a simple annihilation process $\chi\chi \leftrightarrow SM$ we can use Boltzmann equation to describe the density of DM particles n as a function of time t:

$$\frac{dn}{dt} = -3\mathcal{H}n - \langle\sigma v\rangle\left(n^2 - n_{eq}\right) \tag{1.45}$$

with $\mathcal{H}$ being Hubble parameter, $\langle\sigma v\rangle$ the thermal average of the annihilation rate and n_{eq} the density value reached at equilibrium. Figure 1.5 shows the evolution of n as a function of x, the ratio between the DM mass m and the temperature of the Universe T ($x \propto t$). When T was very high, the production and annihilation processes impact was much more intense than the one from expansion and thus the density followed this equilibrium. Once $T \leq m$, pair production reduced as well as annihilation, up until the number of particles ceased to evolve and a relic density was left. The final value for the relic density is inversely proportional to the thermal average of the annihilation rate.

The approximate solution of equation 1.45 yields the estimated DM relic density Ω_χ

$$\Omega_\chi h^2 \sim \frac{3 \cdot 10^{27}\mathrm{cm}^3\mathrm{s}^{-1}}{\langle\sigma v\rangle} \tag{1.46}$$

where the Hubble parameter is defined as $h = H_0/100\ \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. Weak-scale values for the interaction cross-section match this result if the DM candidate has masses in the range GeV-TeV.

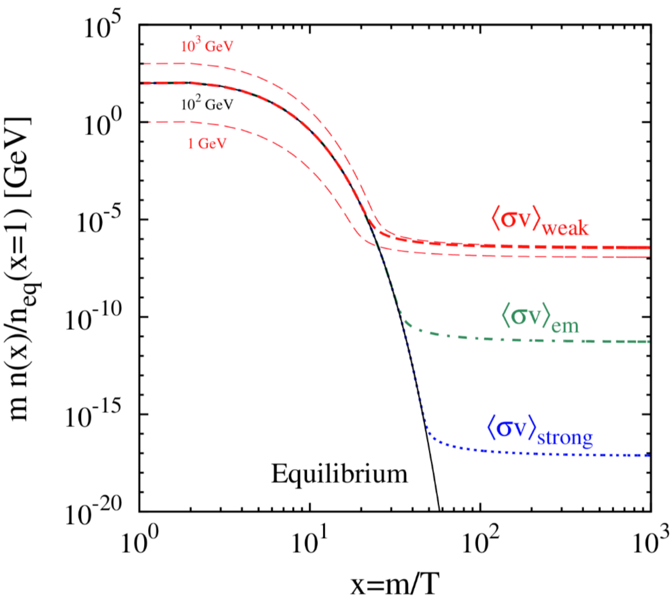


Figure 1.5: Evolution of DM particle number as a function of $x = m/T$. Three different asymptotic limits based on different hypothesis on the cross-section are underlined. [108].

Many theories have been formulated in order to answer these open questions and provide a valid alternative/extension of the SM. The work described in this thesis is about searches for evidence of Physics Beyond Standard Model (BSM) but in a model-independent way, without relying on a specific model prediction, even though the *Heavy Vector Triplet* (HVT) framework [102] has been considered to interpret results and set upper limits on the cross-section.

1.3.1 Heavy Vector Triplet

The HVT model is a simplified parametric Lagrangian, designed to reproduce a large class of explicit descriptions of heavy vector particles (with masses of the order of TeV) in different regions of its parameter space. It provides a Lagrangian that fulfills SM symmetry constraints with an isospin SU(2) triplet formed of a neutral Z' (or V^0) and two charged W' (or $V^\pm$) bosons, defined by the familiar relations:

$$V_\mu^\pm = \frac{V_\mu^1 \mp V_\mu^2}{\sqrt{2}}, \; V_\mu^0 = V_\mu^3 \tag{1.47}$$

The simple phenomenological Lagrangian is the following:

$$
\begin{aligned}
L_V = -\frac{1}{4} D_{[\mu} V^a_{\nu]} D^{[\mu} V^{\nu]a} + \frac{m_V^2}{2} V^a_\mu V^{\mu a} \\
+ i g_V c_H V^a_\mu H^\dagger \tau^a D^\mu H + \frac{g^2}{g_V} c_F V^a_\mu J^{\mu a}_F \\
+ \frac{g_V}{2} c_{VVV} \epsilon_{abc} V^a_\mu V^b_\nu D^{[\mu} V^{\nu]c} + g_V^2 c_{VVHH} V^a_\mu V^{\mu a} H^\dagger H \\
- \frac{g}{2} c_{VVW} \epsilon_{abc} W^{\mu\nu a} V^b_\mu V^c_\nu
\end{aligned}
\tag{1.48}
$$

where $a = 1, 2, 3$. The first line of the equation contains the kinetic and mass term, plus trilinear and quadrilinear interactions with the vector bosons from the covariant derivatives

$$
D_{[\mu} V^a_{\nu]} = D_\mu V^a_\nu - D_\nu V^a_\mu, \;\; D_\mu V^a_\nu = \delta_\mu V^a_\nu + g \epsilon^{abc} W^b_\mu V^c_\nu. \tag{1.49}
$$

The second line contains direct interactions of V with the Higgs current and with the SM left-handed fermionic currents, where $\tau^a = \sigma^a/2$.
The couplings of the HVT to all SM particles are given in terms of the new coupling g_V, which parameterizes the interaction strength between the heavy vectors. This makes the setup extremely versatile since it can capture the features of many, weakly and strongly coupled, concrete models. The relevant parameter space of an HVT with a given mass is two-dimensional consisting of two parameter combinations which describe its couplings to fermions and to SM gauge bosons. The HVT model under consideration is a simplified version with an universal coupling C_F of V to fermions. The coupling of the HVT to fermions scales as $g_F = g^2 c_F / g_V$, where g is the SM $SU(2)_L$ gauge coupling and c_F is a free parameter which can be fixed in each explicit model. In both benchmark models A and B, c_F is expected to be of order one. The HVT coupling to fermions in strongly interacting models is thus g^2/g_V suppressed with respect to weakly coupled models. Thus, in general, a large coupling g_V corresponds to a small Drell-Yan production rate and, similarly, a small branching ratio into fermionic final states. Concerning the HVT coupling to SM bosons, note that it couples dominantly to the longitudinal components of the gauge bosons and to the Higgs, while the coupling to transverse gauge bosons is generally suppressed. Contrary to the coupling to fermions, the HVT coupling to SM bosons scales as $g_H = g_V c_H$. The parameter c_H, analogously to c_F, has to be fixed in each individual model and takes values of order one in models A and B. Here we have the reversed situation, that a small value of g_V in weakly coupled extensions of the SM leads to a small branching fraction into gauge bosons, while strongly coupled theories predict an enhanced branching ratio. Thus strongly and weakly interacting heavy vectors are expected to have a very

different phenomenology: weakly coupled vectors are produced copiously, decay predominantly into two leptons or jets and have a small branching ratio into gauge bosons; strongly interacting vectors are produces less, decay predominantly into gauge bosons and two-fermion final states can be extremely rare. The results can be presented as contours in the parameter space (g_H, g_F) which allow for a broad interpretation, going beyond the benchmark models A and B. In a very large class of explicit models of heavy vectors, the parameters c_H and c_F can be computed and the result compared with the aforementioned contours to asses the compatibility of the concrete model with the experimental search.
The two benchmark models, A and B, are predominately produced via quark-anti-quark annihilation. To study the rare process of vector-boson-fusion a third model, model C, is designed to focus on this production mode. In this model the couplings are set to $g_H = 1$ and $g_F = 0$.
Diboson final states, both neutral W^+W^-, ZH, and charged $W^\pm Z, W^\pm H$, where H is the SM Higgs boson, are particularly interesting in strongly coupled models where the branching ratio into diboson final states is enhanced. Note that the HVT coupling to two SM bosons comes from a gauge invariant coupling to the electroweak triplet Higgs current, with strength g_H, and thus all the couplings to the aforementioned final states are expected to be equal. In particular the HVT framework predicts the same branching ratios for the four processes $V^\pm \to W^\pm Z,\ V^\pm \to W^\pm H,\ V^0 \to W^\pm W^\pm,\ V^0 \to ZH.$
Other neutral diboson final states are either suppressed or forbidden. The decay of a spin-1 vector into HH is forbidden by momentum and angular momentum conservation (and consequently Lorentz invariance). This is true to all orders, so no higher dimensional operator can appear. ZZ is accidentally not present at dimension four. While operators can appear at dimension six and higher, they are highly suppressed and therefore not considered.
This relation is of primary importance in the HVT framework since it allows to gain a higher sensitivity by combining not only neutral and charged channels, but also eventually channels involving the Higgs boson.
Note that the branching ratios of W' and Z' into bosons are the same. The reason is that the HVT W'/Z' both couple to the Higgs current from which the widths and branching ratios into SM gauge bosons are derived.
The HVT also couples to the fermionic current. Therefore, what needs to be compared is the sum of the widths (or equivalently the sum of the BRs) of all quark and lepton final states. For example, the sum of the widths into $\ell\ell$ and $\nu\nu$ is the same as the width into $\ell\nu$. For the quark sector, the mixing has to be taken into account and the sum of all charged quark final states is equal to the sum of all neutral quark final states.

Although the model-independent nature of the analysis described in this thesis, the

Heavy Vector Triplet (HVT) Model A ($g_V = 1$) produced via qq scattering is used as a baseline for the signal interpretation. The WH configuration is modified to replace the W with the X particle. The X is considered to be charged with variable mass, natural width of $2\,\mathrm{GeV}$ (smaller than the detector resolution), and spin-1, with allowed decays only to $u\bar{d}$ with 100% branching ratio.

CHAPTER 2

ATLAS Experiment at LHC

The European Organization for Nuclear Research (CERN) is a research center that runs the world's largest particle physics laboratory, based in the north-west of the city of Geneva on the Franco-Swiss border. Its origins can be traced to the 1940s, when a small number of visionary scientists in Europe and North America identified the need for Europe to have a world-class physics research facility.
At an intergovernmental meeting of UNESCO in Paris in December 1951, the first resolution concerning the establishment of a European council for nuclear research was adopted. Two months later an agreement was signed establishing the provisional Council – the acronym CERN was born. Following the ratification by 12 states (Belgium, Denmark, France, the Federal Republic of Germany, Greece, Italy, the Netherlands, Norway, Sweden, Switzerland, the United Kingdom and Yugoslavia), on 29 September 1954 the European Organization for Nuclear Research officially came into being.
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's main area of research is sub-atomic particle physics and it is run by 23 Member States (Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovak Republic, Spain, Sweden, Switzerland and United Kingdom) while many non-European countries are involved in different ways. Among the most important achievements made by scientists through experiments at CERN we can cite the discovery of the W, Z and Higgs boson [85, 26, 27, 68, 67, 50, 64], the discovery of direct CP violation [65] 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 $7\,\mathrm{TeV}$ 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 $50\,\mathrm{MeV}$ energy. The *Proton Synchroton Booster* and sequentially the *Proton Synchroton* (PS) accelerate the protons to $1.4\,\mathrm{GeV}$ and $25\,\mathrm{GeV}$, respectively. The beam is then injected into the *Super Proton Synchrotron* (SPS) to be accelerated to $450\,\mathrm{GeV}$ and finally to LHC. Most of the accelerators before LHC have their own halls where experiments exploit the beams at lower energies.
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 proton beams enter the LHC divided in bunches of 10^{11} protons each, at a frequency of 40 MHz hence separated in time by $25\,\mathrm{ns}$, travelling in opposite directions in such a way that several collision points are found along the ring. The bunch fill pattern is designed to maximize the rate of the collisions for a total of 2556 proton filled bunches in Run-2, out of a maximum allowed of 2808. Four detectors are installed near the collision points: *ALICE*, *ATLAS*, *CMS* and *LHCb*. The nominal project value of the center of mass energy at each collision point equals $13\,\mathrm{TeV}$, but since the first day of Run-3 the record energy of $13.6\,\mathrm{TeV}$ has been reached.
Protons are not the only particles accelerated in the LHC, also 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.
A construction parameter that is crucial to the discovering power of a collider is its instantaneous luminosity $\mathcal{L}$, which represents the number of collisions that can be

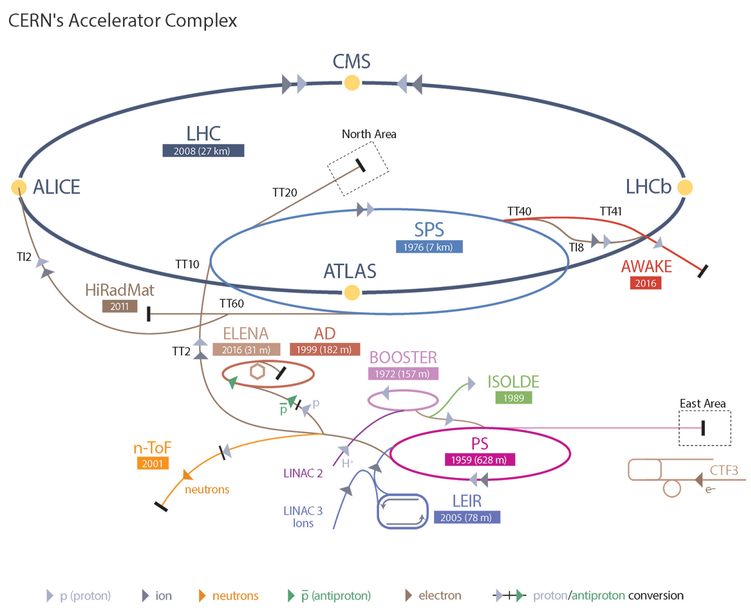


Figure 2.1: Scheme of the CERN accelerator complex, from Linac2 to LHC and a subset of the many experiments supported by these accelerators.

produced in a particle collider per cm^2 and per second:

$$\mathcal{L} = f_{\text{rev}} F \frac{N_b^2 n_b}{2\pi\sigma_x\sigma_y} \tag{2.1}$$

It is directly proportional to the revolution frequency of the bunches f_{rev} the number of protons per bunch N_b and the number of bunches n_b. It is also inversely proportional to the root mean square of the beam width in the x and y directions σ_x and σ_y, and corrected by a geometrical form factor F which accounts for the crossing angle of the two beams. 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:

$$R = \mathcal{L} \cdot \sigma \tag{2.2}$$

The total amount of luminosity delivered by the LHC is called integrated luminosity and is defined as:

$$L = \int \mathcal{L} dt \tag{2.3}$$

It is used to quantify the amount of data delivered by the LHC and recorded by the experiment.
The extreme experimental environment caused by the intense luminosity involves a high particle multiplicity which is particularly challenging both at the detector and trigger system level. This is caused by the so called in-time pile-up, arising from multiple collisions from the same bunch crossing, and out-of-time pile-up, which originates from signals coming from previous bunch. The parameter used to describe this aspect is the average number of interactions per bunch crossing $\langle\mu\rangle$:

$$\langle\mu\rangle = \frac{L \cdot \sigma_{\text{inelastic}}}{n_b \cdot f_{\text{rev}}} \tag{2.4}$$

2.1.1 Run-1

The first successful circulation of the beams occurred on 10^{th} September 2008. Unfortunately some of the superconducting magnets suffered damage, causing the LHC to stop its activity until November 2009, after a long technical intervention. First collisions took place on 30^{th} March 2010, with the rest of the year mainly devoted to commissioning.
2011 was the first production year, with an integrated luminosity of $5\,\text{fb}^{-1}$ delivered to both ATLAS and CMS. By the time of the 2012 summer conferences, data corresponding to $6\,\text{fb}^{-1}$ were collected, allowing for the announcement of the discovery of a Higgs-like particle on 4^{th} July 2012.
During the first period (2010-2011), the centre-of-mass energy of the pp system was $\sqrt{s} = 7\,\text{TeV}$. In 2012, $\sqrt{s}$ was increased to $8\,\text{TeV}$.
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 $1 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. Such a high luminosity was obtained through 1380 bunches per beam with a collision rate of $50\,\text{MHz}$, values smaller than the nominal ones. The resulting 2011-2012 integrated luminosity is shown in Figure 2.2 and corresponds to $21.3\,\text{fb}^{-1}$ of data recorded by the ATLAS detector. Figure 2.3 gives an overview of the luminosity-weighted distribution of $\langle\mu\rangle$ for the Run-1. The mean number of interactions reached the value of 9.1 at $\sqrt{s} = 7\,\text{TeV}$ and 20.7 at $\sqrt{s} = 8\,\text{TeV}$. The LHC Run 1 officially ended in 2013.

2.1.2 Run-2

A phase of long stop for LHC, the Long Shut-down 1 (LS1), lasted from 2013 to 2014 to allow maintenance and upgrade operations. In May 2015 a second phase of data collection started: the LHC Run-2. The goal of Run-2, lasted until the end of 2018, consisted in collecting about $100\,\text{fb}^{-1}$ of data at a collision energy of

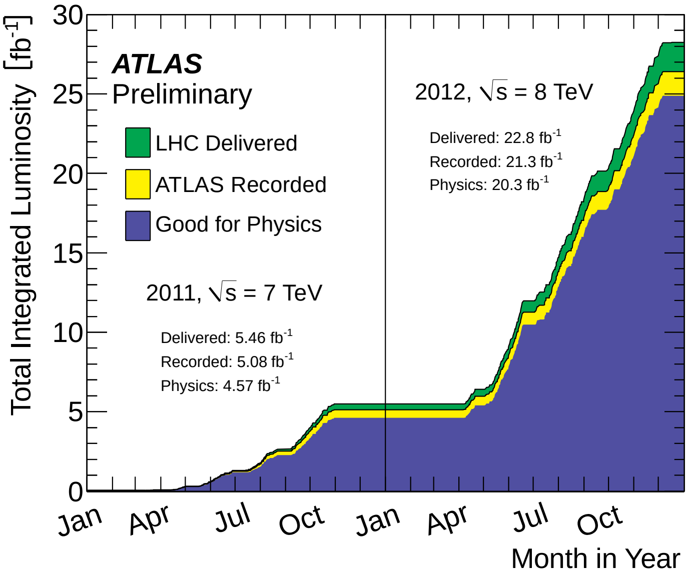


Figure 2.2: Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow) and certified to be good quality data (blue) during stable beams and for pp collisions at 7 and $8\,\mathrm{TeV}$ centre-of-mass energy in 2011 and 2012.

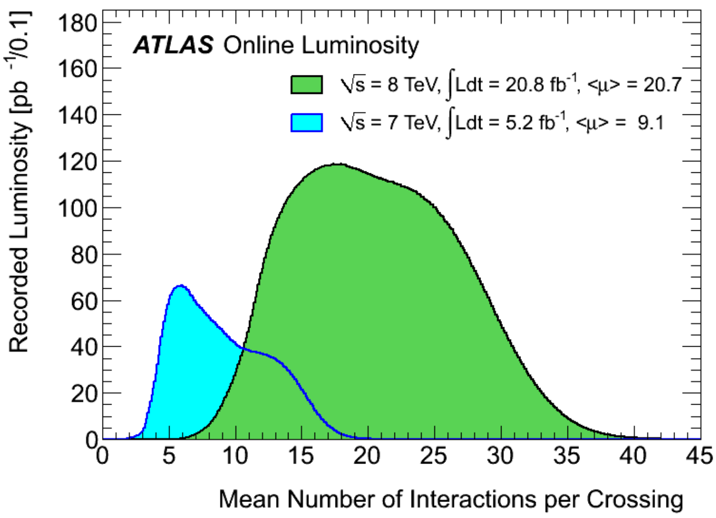


Figure 2.3: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2011 and 2012 data. The mean number of interactions per crossing corresponds the mean of the poisson distribution on the number of interactions per crossing calculated for each bunch.

$14\,\mathrm{TeV}$.
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 $50\,\mathrm{ns}$ for $1\,\mathrm{fb}^{-1}$ integrated luminosity. After the commissioning phase, the bunch-spacing was improved to $25\,\mathrm{ns}$ and the beams produced collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ with a peak luminosity of $5.0 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. After the Run-2 first technical stop in 2016, LHC reached a peak luminosity of $12.1 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. 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 $40\,\mathrm{fb}^{-1}$, compared with the $25\,\mathrm{fb}^{-1}$ 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 $2.09 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.
In the same year the LHC has provided its two major experiments, ATLAS and CMS, with $50\,\mathrm{fb}^{-1}$ 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 $60\,\mathrm{fb}^{-1}$ for the ATLAS and CMS proton run was exceeded by 10%, resulting in a total integrated luminosity during Run-2 (2015 - 2018) of $140\,\mathrm{fb}^{-1}$, and of $139\,\mathrm{fb}^{-1}$ since the start of LHC physics. Figure 2.4 shows the integrated luminosity recorded by the ATLAS experiment during Run-2.
Figure 2.5 gives an overview of the luminosity-weighted distribution of $\langle\mu\rangle$ for the Run-2. The mean number of interactions reached the value of 36.1 in 2018.

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 but due to Covid-19 pandemic it was prolonged till the end of spring of 2022. During this time the four big LHC experiments performed major upgrades to their data readout and selection systems, with new detector systems and computing infrastructure. The changes will allow them to collect significantly larger data samples, of higher quality than previous runs, with plans to start Run-3 and, successively, guide the LHC towards its High

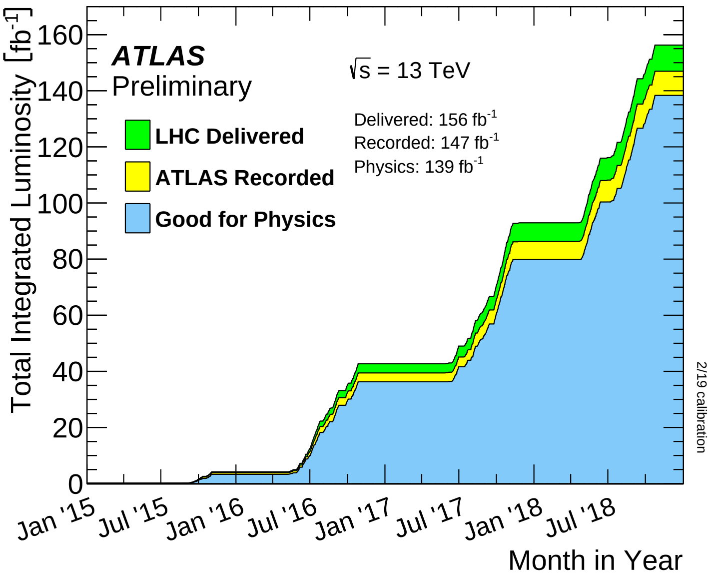


Figure 2.4: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at $13\,\mathrm{TeV}$ centre-of-mass energy in 2015-2018. The complete pp data sample in 2018 is shown.

Luminosity phase.
LHC Run-2 provided the scientific community with relevant precision measurements, from Higgs boson mass to top quark and vector boson masses. Precision measurements are crucial to constrain Standard Model parameters that are not predicted by the model, hence the need to further increase the statistics of collected data.
First stable-beam collisions of Run-3 has occurred on 5^{th} July 2022, breaking a new energy world record of $13.6\,\mathrm{TeV}$ in the centre-of-mass. With the increased data samples and higher collision energy, Run 3 will further expand the already diverse LHC physics program, lasting until the end of 2025, and will act as a test-bench for the new technologies installed in anticipation of *High Luminosity LHC* (HL-LHC).
The HL-LHC is an upgrade of the LHC scheduled to run from 2027 and designed to achieve an instantaneous luminosity 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 program. As for the ATLAS detector, the HL-LHC will deliver a luminosity seven times 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

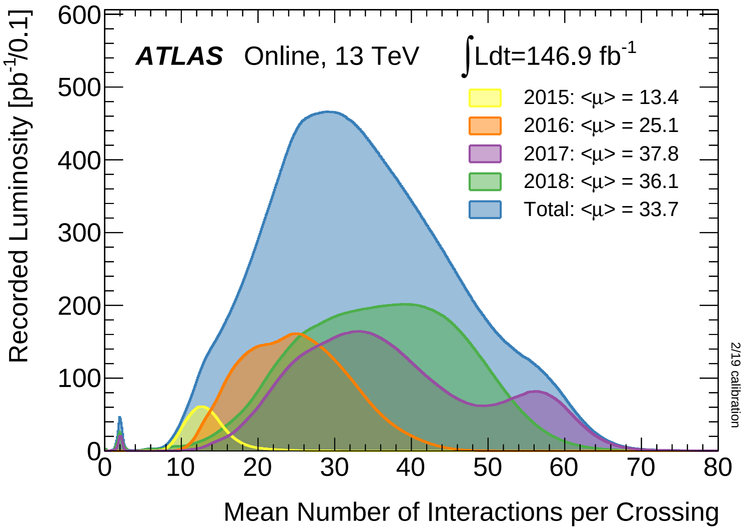


Figure 2.5: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2015 – 2018 pp collision data at 13 TeV centre-of-mass energy. The mean number of interactions per crossing corresponds the mean of the poisson distribution on the number of interactions per crossing calculated for each bunch.

granularity and faster readout are needed. While most of these new systems will be installed during the next Long Shutdown (LS3, scheduled for 2026), some have been already installed in LS2, as the ATLAS New Small Wheels described in Appendix A.

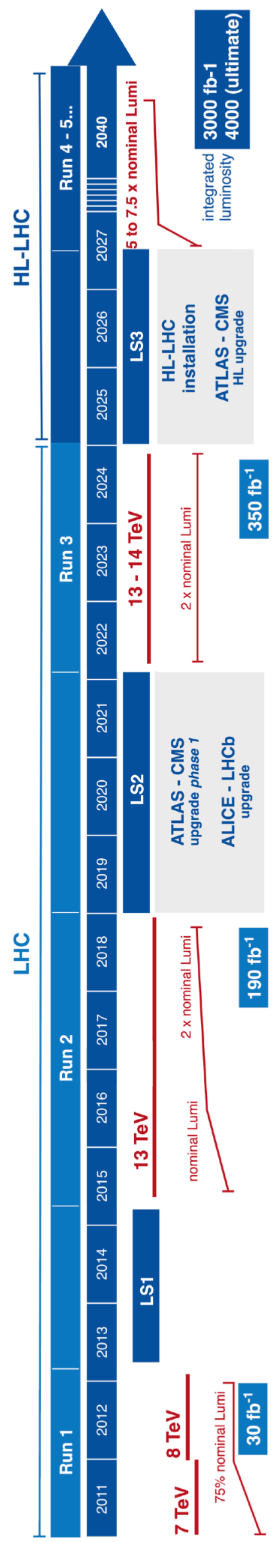
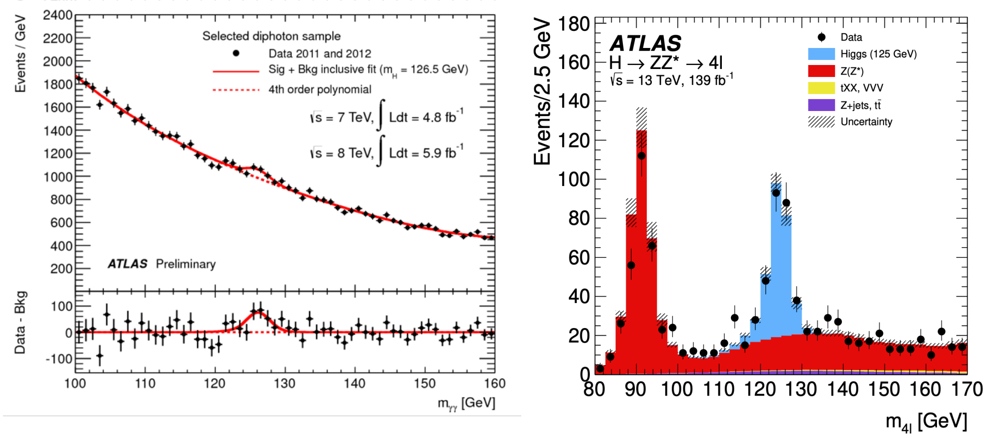


Figure 2.6: LHC timeline from the start of Run-1 to the approximate HL-LHC program.

2.2 The ATLAS detector

The ATLAS experiment is a multi-purpose experiment at the Large Hadron Collider, the biggest ever built, designed to exploit the full discovery potential and the huge range of physics opportunities that LHC provides. Its name is an acronimous of *A Toroidal Lhc ApparatuS*, for the peculiar toroidal form of the magnetic field provided by superconductive magnets. It sits in a cavern situated $100\,\mathrm{m}$ below ground near the main CERN site. Together with CMS, it was one of the two LHC experiments involved in the discovery of the Higgs boson in July 2012 [50, 64] (Figure 2.7), but it was also designed to search for evidence of theories of particle physics beyond the Standard Model.

A schematic representation of the ATLAS detector is shown in Figure 2.8. It



(a) (b)

Figure 2.7: 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 $125\,\mathrm{GeV}$ is consistent with predictions for the Standard Model Higgs boson (a). A recent [55] distribution of candidate Higgs events from the H to ZZ to 4 leptons analysis using $13\,\mathrm{TeV}$ data from the LHC (b).

consists of a cylindrical geometry section around the beam axis, called *barrel*, closed at sides 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.

The subsystems characterizing the detector, from inside to outside are the following:

- 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.

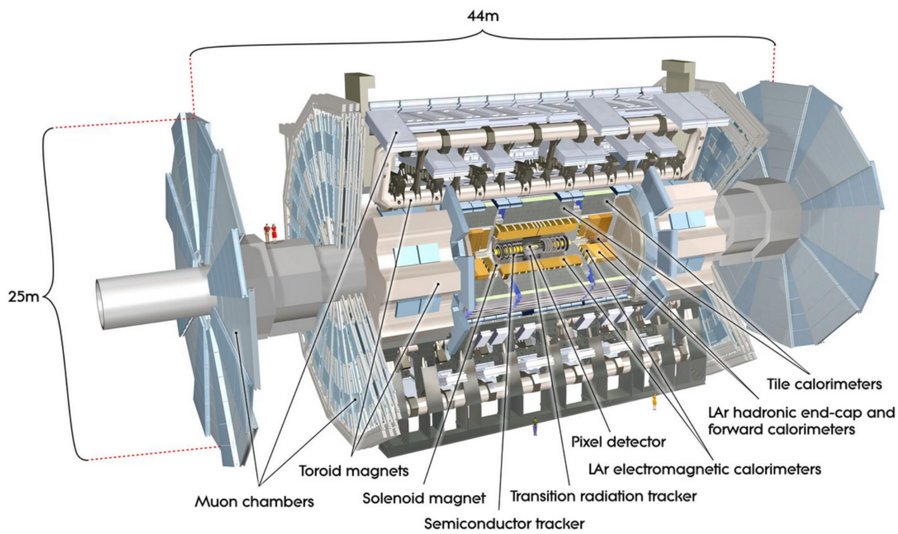


Figure 2.8: A digital reproduction of the ATLAS experiment with its subdetectors.

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 2.9). Due to the symmetry of the detector, a cylindrical coordinates system reference is actually preferred, 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

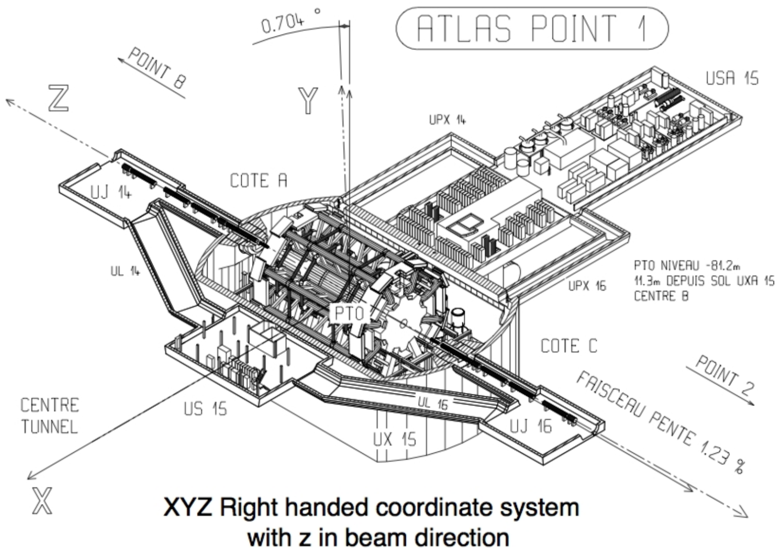


Figure 2.9: Coordinate system used in ATLAS.

boosts along the z axis. In contrast, the *rapidity* variable defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) \tag{2.5}$$

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 (m $\ll$ **p**) $E \approx p$, therefore it can be written as:

$$y = \frac{1}{2} \ln \left(\frac{p + p_L}{p - p_L} \right) = - \ln \left(\sqrt{\frac{(1 - cos\theta)}{(1 + cos\theta)}} \right) == - \ln \left(\sqrt{\frac{sin^2\theta}{(1 + cos\theta)^2}} \right) \tag{2.6}$$

Given the equality $\frac{sin\theta}{(1+cos\theta)} = \tan\left(\frac{\theta}{2}\right)$, the expression becomes the *pseudo-rapidity*:

$$\eta = - \ln \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{2.7}$$

that is used in place of θ given its Lorentz invariance.
It is easy to see that $\eta = 0$ for $\theta = 90°$ and asymptotically increases as $\theta \to 0°$ (or $\theta \to 180°$), as shown in Figure 2.10.

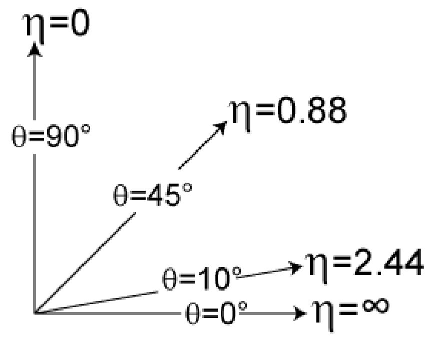


Figure 2.10: Pseudorapidity η for different values of θ angle.

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

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.8}$$

is useful and largely used in physics analyses.

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

2.2.2 The Inner Detector

The first detector encountered moving inside-out from the beam-pipe is the Inner Detector (ID). It is responsible for the tracking of charged particles produced in the pp collisions alongside with their kinematic features: p_T, η, ϕ and the reconstruction of primary and secondary production vertices.
The ID extends from 38 to 1082 mm in radius, and for 6.2 m in length, covering the range in pseudorapidity $|\eta| < 2.5$. It is composed by three subsystems: The Pixel Detector (PD), the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT), as shown in Figure 2.11.

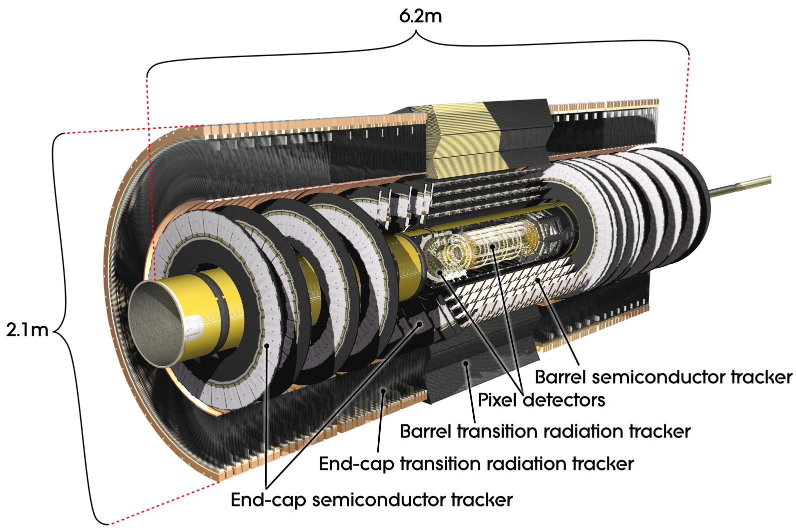


Figure 2.11: The ATLAS Inner Detector subsystems structure.

Pixel Detector

The Pixed Detector [44, 17] is composed of three cylindrical layers at radius 50.5, 88.5 and 122.5 mm in the barrel region and three disc-shaped layers for the end-caps at longitudinal distance of 49.5, 58.0 and 65.0 mm. Each layer is made of pixels arranged in plates of 4086 elements each. Pixels have size $400 \times 50 \mu m^2$ with a resolution of 10 μm in the $R - \phi$ direction and 115 μm in the z direction. In 2014 the so-called Insertable B-layer, a fourth innermost layer with smaller pixels ($50 \times 250\ \mu m^2$), was added at radius 33.3 mm in order to significantly increase the vertex reconstruction capability.

Semiconductor Tracker

The SemiConductor Tracker (SCT) [29] is composed of silicon micro-strip modules layers, four in the barrel region and nine in the end-caps. Each module is composed of 768 read-out strips with a pitch of $80\ \mu m$, and dimensions $6.36 \times 6.40\ cm^2$. They are disposed in a stereo configuration, providing a precision measurement in the principal coordinate R-ϕ of 17 μm, and in the second coordinate z of $580\ \mu m$. The entire system is mapped to more than 6 million read-out channels.

Transition Radiation Tracker

The Transition Radiation Tracker (TRT) [110] constitutes the outer part of the ID. It is a gas detector made of $4\ mm$ diameter tubes, filled with a 70% Xe, 27% CO_2 and 3% O mixture. The ionization electrons produced by particles traversing the detector is collected through a gold-plated tungsten cable at the center of each tube. The TRT is used to reconstruct the tracks and, owing to the so-called transition radiation[1], provides information on the particle type that flew through the detector: it is capable of a significant discrimination between electrons and charged pions with energy in the range $1\,\mathrm{GeV} \leq \mathrm{E} \leq 100\,\mathrm{GeV}$.

Through the track measurement provided by ID it is possible to obtain a measure of the transverse momentum of the charged particle by exploiting the information about the curvature of the track caused by the magnetic field, the latter being provided by the superconducting solenoid described in Section 2.2.5. In fact, a particle of charge q and velocity $\vec{v}$, when surrounded by a magnetic field $\vec{B}$, undergoes the Lorentz force

$$\vec{F}_L = q\vec{v} \times \vec{B} \tag{2.9}$$

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

$$r = \frac{p_T}{0.3 \cdot B} \tag{2.10}$$

if r, B and p_T are expressed in m, T, and GeV, respectively. In order to estimate the resolution, the *sagitta method* is used. 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 (in m), 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

$$\frac{\Delta p}{p^2} = \frac{\epsilon}{0.3 \cdot B \cdot L^2}\sqrt{\frac{720}{N+4}}\,. \tag{2.11}$$

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.

[1]The transition radiation is a form of electromagnetic radiation emitted when a ultra-relativistic charged particle passes through inhomogeneous media, such as a boundary between two different media. It is suitable for particle discrimination since the total energy loss of a charged particle on the transition depends on its Lorentz factor $\gamma = E/mc^2$ and is mostly directed forward peaking at an angle of the order of $1/\gamma$ relative to the particle's path.

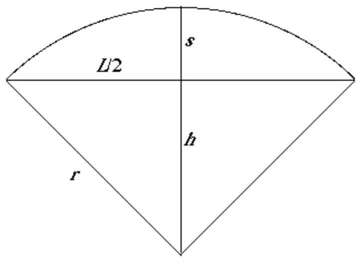


Figure 2.12: Geometric representation of the sagitta.

2.2.3 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 (the only exception being muons and neutrinos), forcing them to deposit all of their energy within the detector. Two types of calorimeters exist: homogeneous and sampling calorimeters, the former being one in which the entire volume is sensitive and contributes a signal, while the latter typically consist in 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. One advantage of this is that each material can be well-suited to its task; for example, a very dense material can be used to produce a shower that evolves quickly in a limited space, even if the material is unsuitable for measuring the energy deposited by the shower. A disadvantage is that some of the energy is deposited in the wrong material and is not measured; thus the total shower energy must be estimated taking into account calibration factors.
ATLAS Calorimeter system is based on sampling method and it consists of four subsystems:

- 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.

A schematic view is given in figure 2.13.

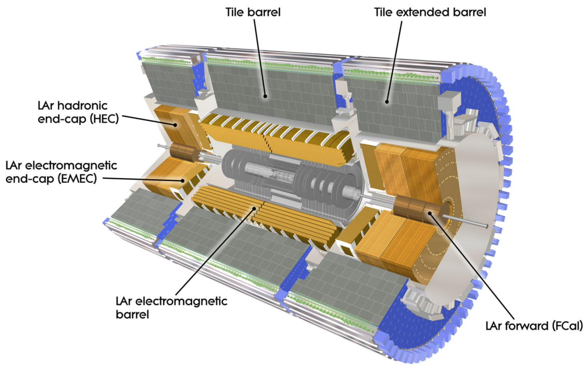


Figure 2.13: The ATLAS Calorimetric system.

Electromagnetic Calorimeter

The measurement of the energy of photons and electrons is based on the production of electromagnetic showers. Therefore the thickness of the detector is typically given in units of X_0, the so-called *radiation length*, indicating the average distance traveled in a certain material before it loses a fraction 1/e of its energy through radiation.
The *Electromagnetic Calorimeter* [21] (ECAL) 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 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. The whole system is maintained at the temperature of 89 K, needed for correct functioning.
The energy resolution of an electromagnetic calorimeter depends on several factors and can be written as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{2.12}$$

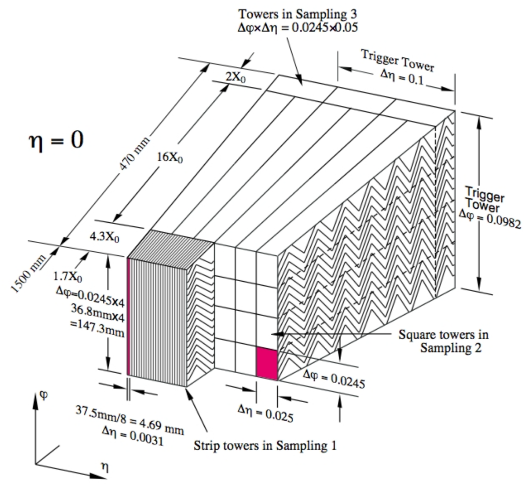


Figure 2.14: Section of ECAL showing its characteristic accordion architecture.

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 $2\,\mathrm{GeV}$ up to $5\,\mathrm{TeV}$.
The angular resolution amounts to about $\frac{40\,\mathrm{mrad}}{\sqrt{E[\mathrm{GeV}]}}$, allowing a good measurement in η of the showers direction.

Hadronic Calorimeters

The hadrons energy loss is usually due to nuclear processes and therefore the spatial scale of hadronic showers is generally larger than the electromagnetic ones. The thickness of the detector is measured in units of interaction lengths λ_i, the mean distance traveled by a hadronic particle before undergoing an inelastic nuclear interaction. 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 λ, in order to guarantee a good resolution for high-energy jets and a good measurement of missing transverse energy, E_T^{miss}. Typically in tens of interaction lengths hadrons are fully absorbed.

The hadronic calorimeters (HCAL), which cover the $|\eta| < 4.9$ pseudo-rapidity range, is divided into three main regions:

- *The Hadronic Tile Calorimeter* [2] is located in the barrel covering the region

$|\eta| \leq 1.0$, and in two extension regions $0.8 \leq |\eta| \leq 1.7$. Steel is used as passive material and scintillating tiles as active materials, which produce a signal proportional to the number of secondary particles produced in the interaction.

- *The Hadronic End-Cap Calorimeter (HEC)* [28] is composed of two independent wheels of radius 2.03 m and covers the range $1.5 \leq |\eta| \leq 3.2$. LAr is used as active medium and copper plates as absorbers, since in the end-caps the amount of radiation is greater than in the barrel.
- *The Forward Hadronic Calorimeter (FCAL)* is placed in the high-eta region very close to the beam pipe where the particle density is extremely high, to cover the region $3.1 \leq |\eta| \leq 4.9$. It is composed of three layers using LAr as active material and copper, for the innermost layer, and tungsten, for the external layers, as abosrbers.

The energy resolution for a hadronic calorimeter can be expressed as

$$\frac{\sigma_E}{E} = \sqrt{\frac{c_{\text{int}}^2 + c_{\text{samp}}^2}{E}} \oplus a \tag{2.13}$$

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.
TileCalo and HEC have an energy resolution of $\frac{\Delta E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$ while FCAL energy resolution is $\frac{\Delta E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$.

2.2.4 The Muon Spectrometer

The more external sub-detector is the Muon Spectrometer (MS) [3], which has the purpose to give a precise and independent measurement of muon quantities, since muons usually pass through the Inner Detector and Calorimeter undetected and reach the outer layer of ATLAS.
The MS 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 spectrometer occupies both the *barrel* and the *end-cap* regions, so that a remarkable pseudo-rapidity coverage in the $|\eta| < 3$ range is guaranteed. It is immersed in a toroidal magnetic field, which allows a muon transverse momentum measurement independent from the one of the ID.
The subdetectors composing the MS, before the arrangement changing for Run-3,

were trigger chambers (Resistive Plate Chambers (RPCs), Thin Gap Chambers (TGCs)), high precision tracking chambers (Muon Drift Tubes (MDTs) and Cathode Strips Chambers (CSCs)). During the LS2 the two Small Wheels in the end-cap region have been replaced by the New Small Wheels, with new kind of detectors more efficient at high collisions rate, with a very small spatial and timing resolution: small Thin Gap Chambers (sTGCs) and MicroMeGas (MMGs), described in Appendix A.
The configuration here described refers to Run-2.

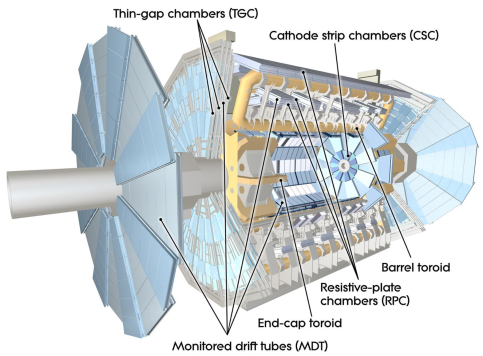


Figure 2.15: The ATLAS Muon Spectrometer, composed of different types of detectors.

In the r - ϕ 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 $5\,\mathrm{m}$, $7.5\,\mathrm{m}$ and $10\,\mathrm{m}$, 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 $7\,\mathrm{m}$, $10\,\mathrm{m}$ and $21\,\mathrm{m}$, respectively.

Resistive Plate Chambers

RPC are gaseous detectors consisting of two bakelite plates separated by a 2 mm gap filled with a gas mixture of 97% tetrafluoroethane ($C_2H_2F_4$) and 3% isobutane (C_4H_{10}). A high electric field of 4.5 kV/mm is maintained between the two plates to amplify the primary ionization of charged particles crossing the detector. The charged induced on metallic strips in the outer sides of the bakelite plates is

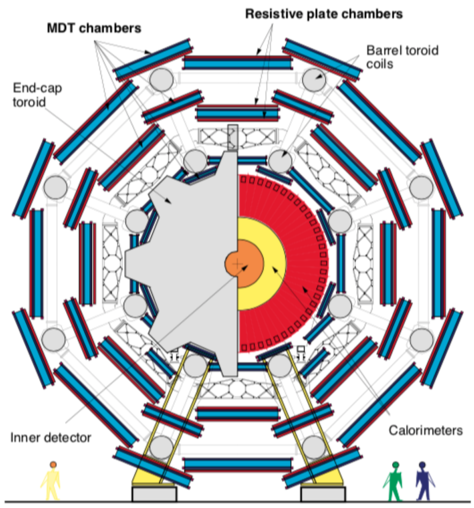


Figure 2.16: The ATLAS Muon Spectrometer section, in the x - y plane.

collected as signal. Two RPC units are placed in each layer, orthogonal to one another, providing information on both η and ϕ coordinates. RPCs are arranged in three trigger station (two in the BM, one in the BO) , each composed by two such layers.

Thin Gap Chambers

TGCs are the trigger chambers used in the end-caps. They are designed to have the anode-cathode spacing smaller than the anode-anode spacing for a very short drift time of 20 ns. The chambers are filled with a highly quenching gas mixture of 55% CO_2 and 45% n-pentane (n-C_5H_{12}) operating in saturation mode.

Monitored Drift Tubes

MDTs are the high precision measurement detectors. They are composed of two multi-layers of aluminum drift tubes, with diameter of 30 mm and thickness of 400 μm, filled with a gas mixture of $93:7-$Ar:CO_2 at a pressure of 3 bar. In the middle of each tube there is a Tungsten-Rhenium wire of 50 μm in diameter. Each chamber typically provides six to eight η measurements along the muon track with a single hit resolution in the precision (rz bending) plane of about 80 μm. The total chamber resolution is 35 μm. The time measurement of the single hit is known with a resolution dictated by the maximum drift time (500 ns). The precision is improved to few nanoseconds when combining hits and fitting to a track.

Cathode Strips Chambers

CSCs consist of multi-wire proportional chambers with cathodes segmented into orthogonal strips. They replace the MDTs that worsen their performance at rates greater than 150 Hz/cm^2 covering the η-region from 2.0 to 2.7. The strips in the transverse plane and parallel to the wires measure the coordinate η and ϕ, respectively. The drift time achieved is less than 40 ns and the precision with which is measured is 7 ns. The spatial resolution reached by CSCs is 40 μm in the radial direction and 5 mm in the second coordinate ϕ.

2.2.5 ATLAS magnets

The ATLAS superconducting magnetic system is $26\,\mathrm{m}$ long, has a diameter of $20\,\mathrm{m}$ 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 2.17).
The Central Solenoid, with an internal diameter of $2.4\,\mathrm{m}$, covers the central region of the detector and provides an uniform magnetic field of approximately $2\,\mathrm{T}$ along the z axis, with a peak of $2.6\,\mathrm{T}$ 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 $3.9\,\mathrm{T}$ for the BT and $4.1\,\mathrm{T}$ 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.

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 $40\,\mathrm{MHz}$.
Each event occupies $\approx 1.5\,\mathrm{MB}$ so that a recording rate of $\approx 60\,\mathrm{TB}$ per second would be needed to save all information from the ATLAS detector, while the current technology allows for a recording rate of $300\,\mathrm{MB/s}$ only. Moreover, it is also true that great part of the collisions lead to relatively non-interesting events for the ATLAS physics program, as the rate is dominated by low-p_T inelastic and diffractive collisions.

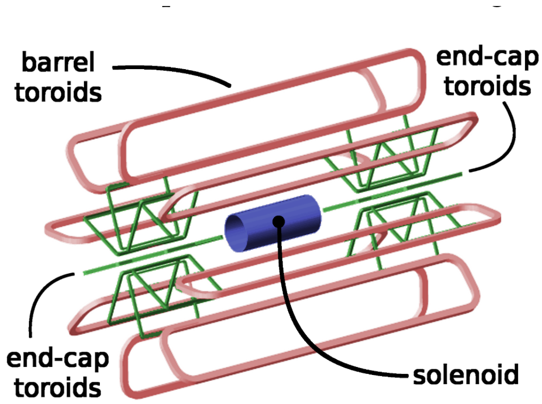


Figure 2.17: Schematic illustration of the magnetic system of the ATLAS detector: in blue the ID solenoid, in green the end-cap toroids and in red the barrel toroid.

The trigger and data acquisition (TDAQ) of the experiment has for these reason the fundamental role of selecting the events which are ”worth” to be recorded, finding the right balance between resources economization and performance. This is a crucial step for everything the collected data is used for, since a non recorded event is lost forever.
The ATLAS Run-2 TDAQ system [99] was built on two levels of online selection, as illustrated in Figure 2.18: a first hardware Level-1 (L1), that significantly reduces the event rate, and a second software-based High Level Trigger (HLT), where the final decision is made.
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.

2.3.1 Level-1 trigger

The L1 trigger is a hardware-based system that uses custom electronics to trigger on reduced granularity information from the calorimeter and muon detectors. The L1 calorimeter (L1Calo) trigger takes signals from the calorimeter detectors as input, providing information about clusters of energy deposits, missing transverse energy and raw shape dimensions using a reduced granularity. The L1 muon (L1Muon) trigger uses hits from the RPCs (in the barrel) and TGCs (in the endcaps) to determine the deviation of the hit pattern from that of a muon with infinite momentum. The hardware L1 exploits quickly accessible coarse data from the calorimeters (L1Calo) and muon systems (L1Muon) in dedicated Regions of Interest (RoI). Information from L1Calo and L1Muon are sent to the Central Trigger Processor (CTP) which decides if to accept or reject the event. At L1 the

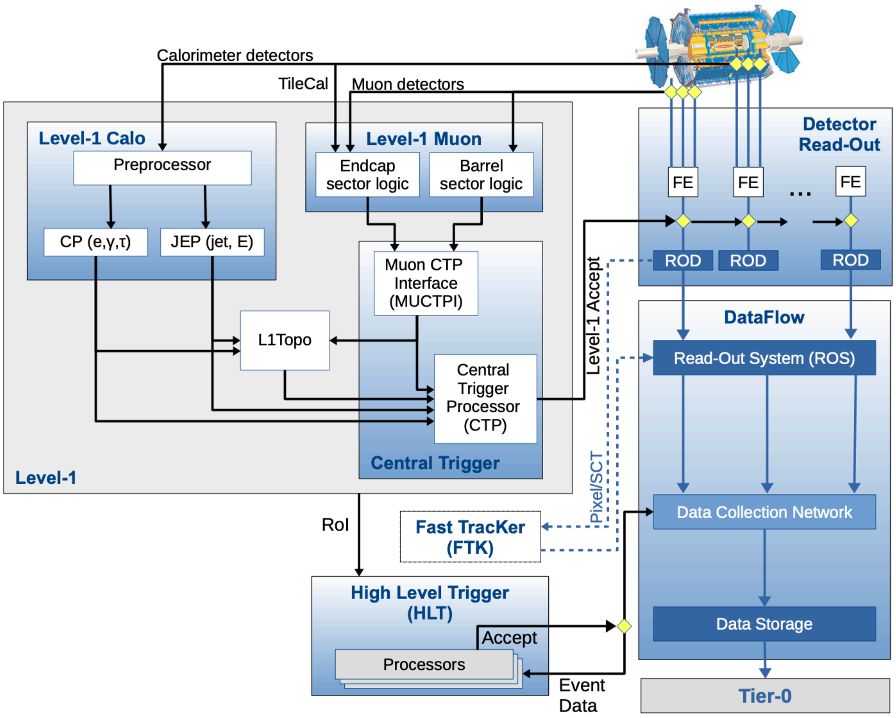


Figure 2.18: Schematic view of the ATLAS Run 2 TDAQ system. [66]

event rate of 40 MHz is reduced to 100 kHz with a 2.5 μs latency.

2.3.2 High Level Trigger

The second stage of the trigger, the HLT, is software-based. It integrates the RoI data with the full detector information and runs complex trigger algorithms to select the events. A fast reconstruction step is first used for the trigger selection, followed by a more precise refinement similar to the offline reconstruction. The HLT is the first step in which ID information is incorporated in the trigger, using only track information inside the identified (η, ϕ) RoI at L1 due processing time constraint. The muon fast reconstruction integrates each L1 muon candidate with MDT data preforming a track fit extrapolated to the ID. The ID fast tracking consists in trigger specific pattern algorithms, designed to identify compatible track segments and hit points. Raw calorimetric informations are reconstructed by fast algorithms into cluster and cell objects, which will later be reconstructed in jet, electron, and photon candidates. The final event rate is reduced to 1kHz with a latency of milliseconds.

CHAPTER 3

Physics Object definition

Physics objects are the building bricks of every ATLAS analysis, and their definition can have a huge impact on the final results. The outputs of the detectors, generated both by real collisions and simulated ones, are processed by a series of off-line algorithms to reconstruct leptons, jets, photons or missing transverse energy, whom four-momentum is used to obtain the final physics result.
In this chapter an overview of the identification and reconstruction algorithms used for the most relevant physics objects and their relative performances is presented. Neither leptons or photons are present in the analysis discussed in this thesis but for the completeness of the speech they are here described.

3.1 Electrons reconstruction

Electrons inside ATLAS leave their traces within the inner detector and the electromagnetic calorimeter.
In the barrel region ($|\eta| < 2.47$) both detector are used while in the forward region only measurements from the EM calorimeter are adopted to reconstruct electrons. The full reconstruction process proceeds by the following steps:

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

Let's start from the reconstruction. An electron from the interaction point loses a significant amount of energy due to bremsstrahlung as it passes through the detector

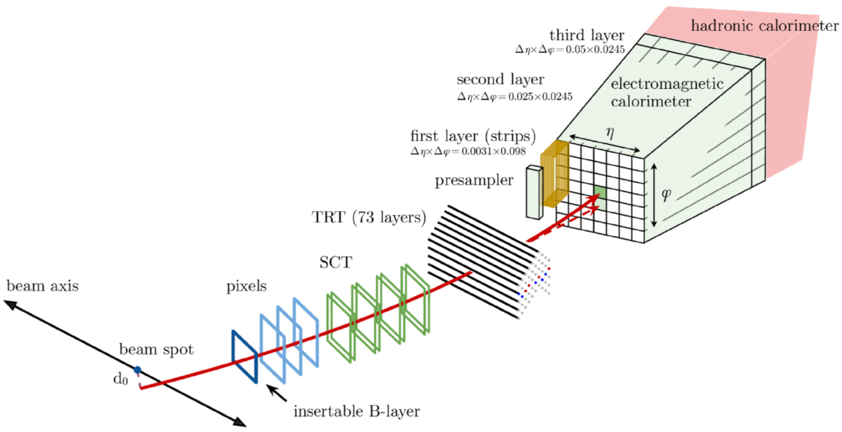


Figure 3.1: A schematic illustration of electrons path through the detector. The red trajectory shows the hypothetical path of an electron, while the dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system.

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. Figure 3.1 displays a schematic picture of the elements involved with the reconstruction and identification of electrons. Electron reconstruction in the barrel region 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 [93] of size 3×5 towers in $\eta - \phi$, whose summed transverse energy exceeds $2.5\,\mathrm{GeV}$. 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 [106]. Track candidates with $p_T > 400\,\mathrm{MeV}$ are fit, according to the hypothesis used in the pattern recognition, using the ATLAS Global χ^2 Track Fitter [69]. 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) [5] 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 [77] 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 $1.37 < |\eta| < 1.52$.

3.1.1 Identification and isolation

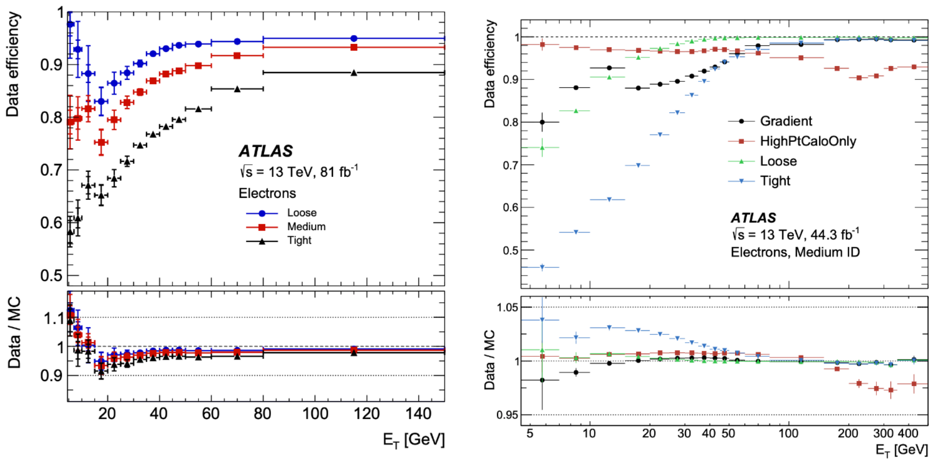
Identification of leptons is carried out with a likelihood-based approach. A Likelihood function is constructed including information from the tracking system

or the calorimeter system and quantities that combine both of them [53], where correlations are neglected.
Here the signals are considered to be the prompt electrons and 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:

$$d'_L = -\tau^{-1} \ln\left(d_L^{-1} - 1\right) \qquad \text{with} \qquad d_L = \frac{L_S}{L_S + L_B}\,. \tag{3.1}$$

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 likelihood-based 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 discriminant. As an example, the efficiencies for identifying a prompt electron with $E_T = 40\,\mathrm{GeV}$ 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 3.2a.
To suppress the background due to non-prompt leptons, e.g. from decays of hadrons produced in jets, the leptons in the event are usually required to be isolated. A calorimeter isolation, a track isolation or both can be applied. The calorimeter isolation is estimated using the energy in a cone of $R = 0.2$ centred around the electron after the subtraction of the energy associated with the electron itself (`EtCone20`). Track isolation is calculated using the scalar sum of tracks p_T in a cone of $\Delta R = 0.3$ centred around the electron without including the electron p_T itself (`PtCone30`). The calorimeter isolation variables usually include a correction for the increase in the energy of the electron in the isolation cone with electron p_T (transverse shower leakage) and for additional energy deposits from pile-up events.

Figure 3.2b shows the isolation efficiencies measured in data and the corresponding data-to-simulation ratios as functions of the electron E_T and η for several operating points and for candidate electrons satisfying Tight identification requirements. The efficiencies that determine the values of the requirements given in Table **??** are evaluated in simulation from a $J/\Psi \to ee$ sample for $E_T < 15\,\mathrm{GeV}$ and from a $Z \to ee$ for $E_T > 15\,\mathrm{GeV}$. The performances of the reconstruction, identification and isolation algorithms are evaluated both in the data and in the MC simulation using the electrons coming from the two resonant processes $Z \to ee$ and $J/\Psi \to ee$ [53].



(a) (b)

Figure 3.2: (a): Measured electron-identification efficiencies in $Z \to ee$ events for the Loose (blue circle), Medium (red square) and Tight (black triangle) operating points as a function of E_T. The bottom panels shows the data-to-simulation ratios. (b): isolation efficiencies for data (in the upper panels) and the ratio to simulation (lower panels) for several operating points as a function of candidate-electron E_T. [14]

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.

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.

The reconstruction proceeds according to five main reconstruction strategies, leading to the corresponding muon types [11]:

- *Combined (CB)* are reconstructed matching tracks from the ID to tracks from the MS and performing a combined fit of all the hits associated.

- *Inside-out combined (IO)*; their reconstruction starts from ID tracks which are then associated to at least three loosely-aligned hits in the MS. ID and MS hits are then fitted together with energy losses in the calorimeter. Since no independent reconstruction in the MS is required, some efficiency is recovered.

- *Muon-spectrometer extrapolated (ME)* use information from the MS only, allowing to exploit its full geometrical coverage, up to $|\eta| < 2.7$.

- *Segment-tagged (ST)* are identified by requiring that an ID track extrapolated to the MS satisfies tight angular matching requirements to at least one reconstructed MS segment. A successfully-matched ID track is identified as a muon candidate, and the muon parameters are taken directly from the ID track fit.

- *Calorimeter-tagged (CT)* exploit deposits in the calorimeter compatible with minimum ionizing particles and found by extrapolating ID tracks through it to tag the candidate as muons. The muons parameters are then extracted from corresponding ID tracks.

3.2.1 Identification and isolation

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': 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.

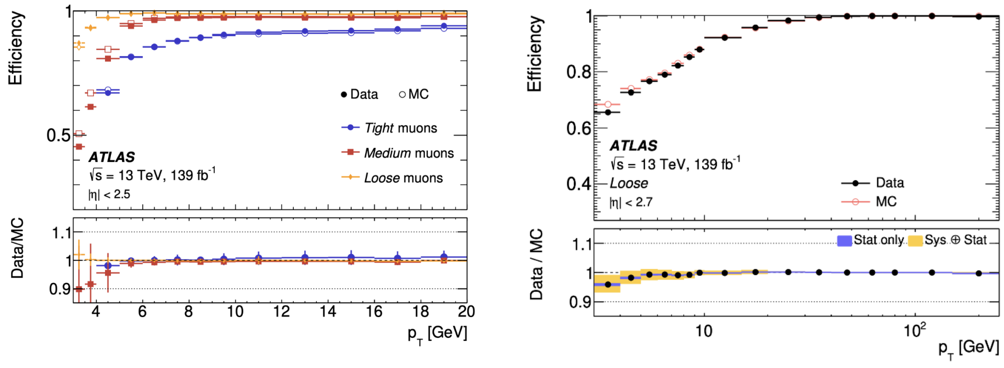
To guarantee a robust momentum measurement, specific requirements on the number of hits in the ID and MS must be met. Four different ”muon” definitions are employed: *Medium*, *Loose*, *Tight* and *High-p_T*.

- *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%.

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 $p_T > 1\,\text{GeV}$ in a cone of size $\Delta R = \min(10\,\text{GeV} \times p_T^{\mu}, 0.3)$ around the muon of transverse momentum p_T^{μ}, excluding the muon track itself. The calorimeter-based isolation variable, $E_T^{\text{varcone30}}$, is defined as the sum of the transverse energy of topological clusters [51] 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.
The three working points are efficiencies as evaluated on a $J/\Psi \to \mu\mu$ sample and are shown in Figure 3.3a. Isolation working points are defined as well, using track- and calorimeter-based isolation variables analogues to the ones defined for electrons; the efficiency for the Loose working point as evaluated on a $Z \to \mu\mu$ sample are shown in Figure 3.3b.



(a) (b)

Figure 3.3: On the left, muon reconstruction and identification efficiencies measured in $J/\Psi \to \mu\mu$ events as a function of p_T for the Loose, Medium, and Tight criteria. The predicted efficiencies are depicted as open markers, while filled markers illustrate the result of the measurement in collision data. On the right side, muon isolation efficiency measured in $Z \to \mu\mu$ events for the Loose working point. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties. [11]

3.3 Photons

3.3.1 Reconstruction

The photons identification starts from the clustering of deposits in the EM and Hadronic calorimeters, called topoclusters [31, 33]. Each topocluster is seeded by a single calorimetric cell with a significance $s = E/\sigma(E)$ grater than 4, and grown by including all adjacent cells with $s > 2$ plus the neighbors having any deposited energy. If two or more local maxima with a deposit greater then $500\,\mathrm{MeV}$ are found, the cluster is split accordingly. The pre-sampling and the first EMC layers are excluded for the initial seeding to suppress the formation of noise clusters.
Each cluster is afterward matched to a fitted track in the ID based on position, taking into account loss due to bremsstrahlung radiation in the tracker. If more than a track is matched to the same topocluster, the one with the most hits in the ID is selected. A converted photon is a cluster matched to a conversion vertex (or vertices), and an unconverted photon is a cluster matched to neither an electron track nor a conversion vertex. About 20% of photons at low $|\eta|$ convert in the ID, and up to about 65% convert at $|\eta| \simeq 2.3$.
Superclusters are then reconstructed adding satellite clusters to the supercluster seed candidates if they satisfy the necessary selection criteria.

The list of clusters is ordered by their transverse energy E_T, obtained as $E_{\text{cluster}}/\cosh(\eta)$. A cluster is elected as photon supercluster seed if it has $E_T > 1.5\,\text{GeV}$ and has not already been used as satellite cluster, with no other requirements on track or conversion vertex matching. Whenever a seed cluster is encountered, satellite clusters are defined and added to the supercluster if they fall within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre, as these cases tend to represent secondary EM showers originating from the same initial electron or photon. A satellite is only added though if its best-matched (electron) track belongs to the conversion vertex matched to the seed cluster. For photons with conversion vertices made up only of tracks containing silicon hits, a cluster is added as a satellite if its best-matched (electron) track belongs to the conversion vertex matched to the seed cluster. These steps rely on tracking information to discriminate distant radiative photons or conversion electrons from pile-up noise or other unrelated clusters.

3.3.2 Identification and isolation

Identification and Isolation are used to select prompt isolated photon over a background from hadronic jets.
Several high-level shape-related shower variables are deployed to define three working points for photons identification: *loose*, *medium* and *tight*. The loose and medium are cut-based working points, while the tight one is optimized through the use of an MVA (Multi-Variate Analysis) and is E_T-dependent. In figure 3.4 the efficiencies of the tight working point are estimated over a $Z \to ll\gamma$ sample as a function of E_T for both converted and unconverted photons using three different methods [33]. The data/MC scale factors are also shown in the lower panel, separately for each measurement and combined.
Photons from prompt decays are usually characterized by low activity around them. Calorimeter and track-based isolation variables are defined to quantify the isolation of photon candidates:

- $\mathbf{E_T}^{coneXX}$: the calorimeter isolation variable, represent the transverse energy of positive-energy topological clusters whose barycentre falls within a cone of $\Delta R = XX$ centered around the photon cluster barycentre after the object energy subtraction and pile-up and calibration corrections.
- $\mathbf{p_T}^{coneXX}$: The track isolation variable is computed by summing the transverse momentum of selected tracks within a cone of $\Delta R = XX$ centred around the photon cluster direction.

With these two variables three working points are defined and their performances are shown in Figure 3.5.

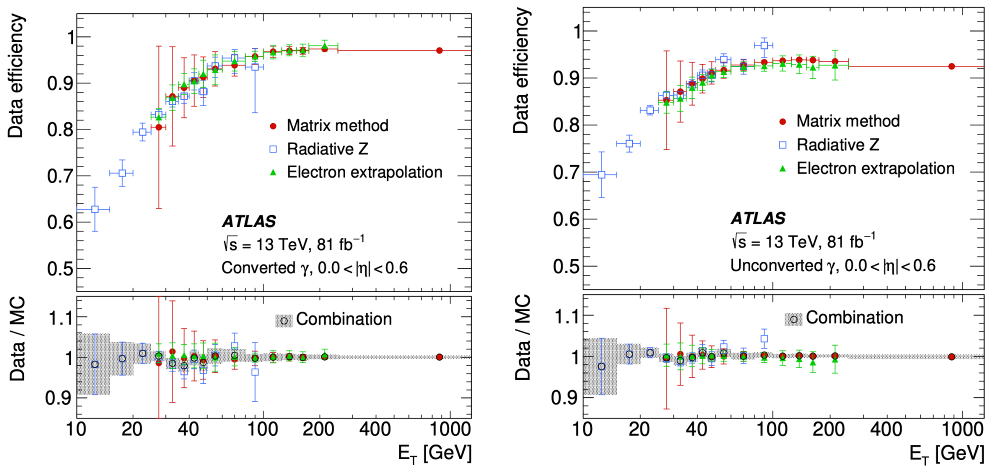


Figure 3.4: The photon identification efficiency, and the ratio of data to MC efficiencies, using $Z \to ll\gamma$ events, for converted (left) and unconverted (right) photons with a Loose isolation requirement applied as preselection, as a function of E_T in the region $0 < |\eta| < 0.6$. The combined scale factor, obtained using a weighted average of scale factors from the individual measurements, is also presented; the band represents the total uncertainty. [14]

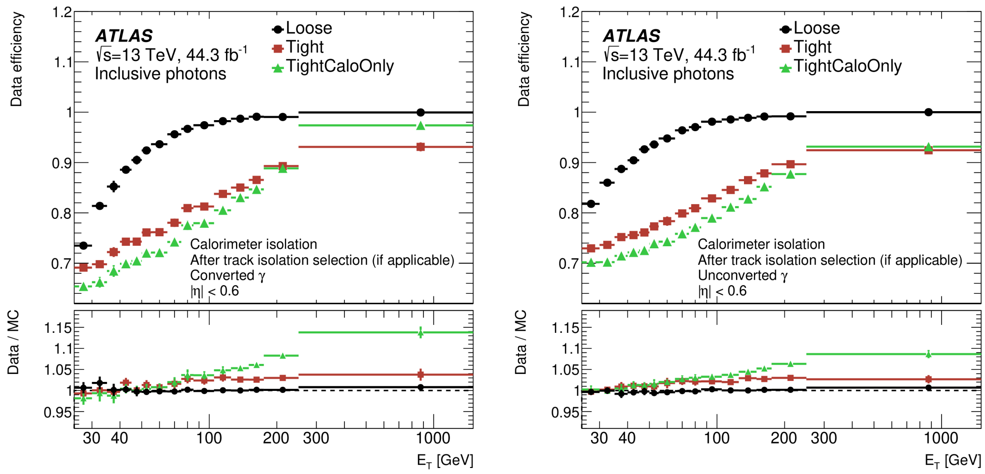


Figure 3.5: Efficiency of the photon isolation working points, using $Z \to ll\gamma$ events, for converted (left) and unconverted (right) photons. The lower panel shows the ratio of the efficiencies measured in data and in simulation. The total uncertainties are shown, including the statistical and systematic components. [14]

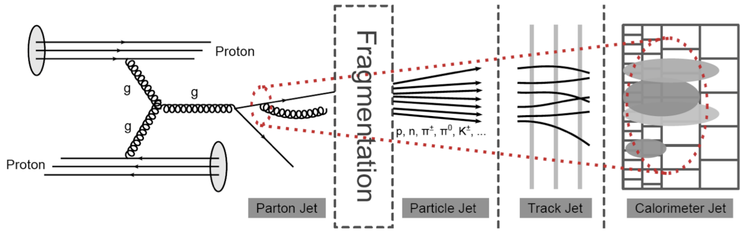


Figure 3.6: Colorful particles, produced at the interaction point, create a bunch of colorless particles due to fragmentation. These particles will produce detector signals in the Inner Detector and Calorimeter, which can be reconstructed as track and calorimeter jets respectively.

3.4 Jets reconstruction

After high energy collisions in a particle collider free quarks or gluons are created. Due to colour confinement, these cannot exist individually. Hadronization is the process of the formation of hadrons out of quarks and gluons and the conic particles burst resulting from this process is called *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 via strong interaction. 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 $10\,\mathrm{ps}$) 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.

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 [52]. 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 [54].
The algorithm used for jets reconstruction is the *anti-k_t* [43] clustering algorithm. It takes four-vector objects as inputs, such as stable particles defined by MC generators, charged-particle tracks, calorimeter energy deposits, or algorithmic combinations of the latter two, as in the case of the particle-flow reconstruction technique [57].
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

$$d_{ij} = \min\left(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,i}^2}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{3.2}$$

where

$$\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2} \tag{3.3}$$

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.

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, while a higher dimension $R = 1.0$ (large-R jets) is implemented when reconstructing the collimated decay products of boosted weak bosons.

In the following sections, an overview of the jets collection adopted in the work described in this Thesis is given.

3.4.1 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 the analysis described in this thesis are reconstructed by means of the particle-flow (*PFlow*) technique [8].

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. These resulting 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 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 $p_T > 500\,\mathrm{MeV}$, 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 [42], 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 [54].

Calibration

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 $e/h > 1$, since the hadronic component of the shower is such that some of the energy deposited is "invisible"[1];

[1]The 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

- 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 [20, 16, 4, 49] 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 $Z \to ee$ 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 below.

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 [54]:

- 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

event-to-event fluctuations.

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 the 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).

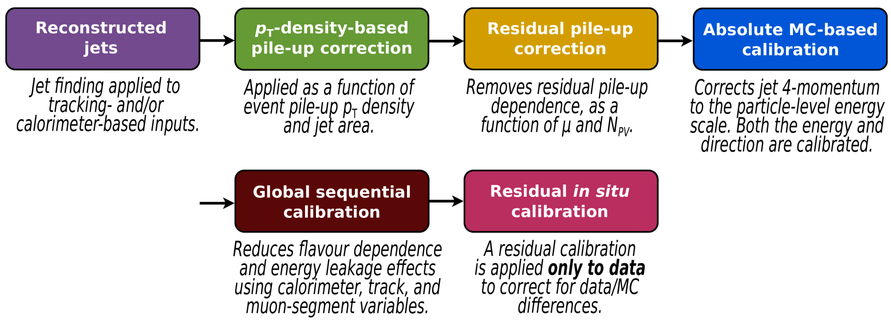


Figure 3.7: Stages of jet energy scale calibrations. Each one is applied to the four-momentum of the jet.

The jet energy resolution (JER) can be parametrized as

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E(\mathrm{GeV})}} \oplus \frac{b}{E(\mathrm{GeV})} \oplus c \tag{3.4}$$

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 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

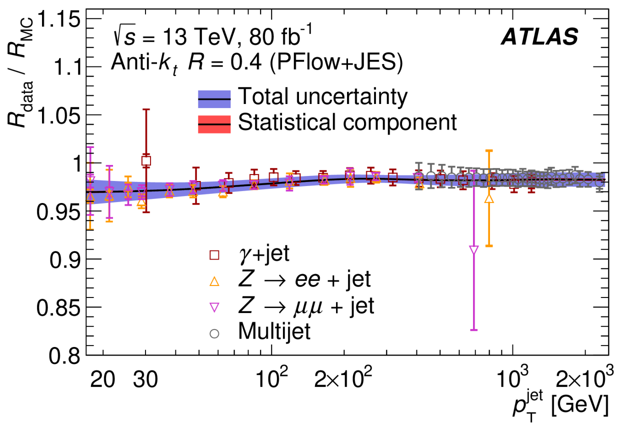


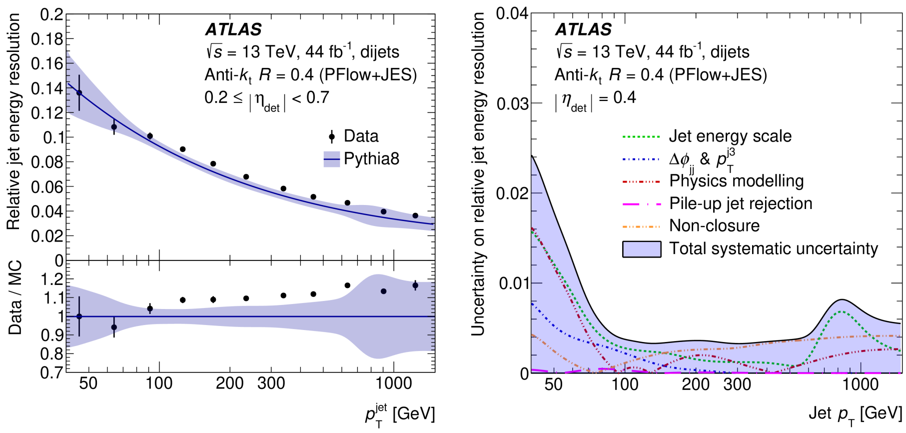
Figure 3.8: Data-to-MC ratio of the PFlow+JES jet response as a function of jet p_T for Z+jet, γ+jet, and multijet *in situ* calibrations. The inner horizontal ticks in the error bars give the size of the statistical uncertainty while the outer horizontal ticks indicate the total uncertainty (statistical 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. [56]

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 [54].

3.4.2 Track-jets

Track jets are jets reconstructed using only the inner detector tracks with the anti-kt algorithm using a distance parameter R $\in \{0.4, 0.3, 0.2\}$. The tracks are required to have $p_T >$ 0.5 GeV, to have at least 1 hit in the pixel detector and 6 hits in the silicon strip detector, and to be tightly matched to the hard scatter vertex using impact parameter thresholds on the tracks. Such thresholds greatly reduce the number of tracks from pileup vertices whilst being highly efficient for tracks from the hard scatter vertex. Once the track jet axis is determined, a second step of track association is performed to select tracks with looser impact parameter requirements in order to collect the tracks needed for effectively running the b-tagging algorithms (see 3.4.4).

An additional jet collection, called *track-jets* [91], clustered with variable ΔR cone sizes by considering only the track information from the ID detector, is de-



(a) (b)

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.[56]

fined. To see how this could be useful, consider a resonance at rest in the lab frame decaying into two partons. To first approximation, the shower of hadrons resulting from each parton will fall in a circular cone of fixed angular size regardless of the orientation of the decay with respect to the beam axis. However, a jet algorithm which uses a fixed cone size in (η, ϕ) will not reflect this behavior, as fixed ΔR corresponds to variable angular size. This can be remedied by letting the cone size of a jet vary as

$$\Delta R = \frac{\rho}{p_T} \tag{3.5}$$

where ρ is a dimensionful constant.

3.4.3 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:

$$\Delta R \sim \frac{2m}{p_T} \tag{3.6}$$

If the momentum is high enough, all the hadronic decay products may be reconstructed as a single large-radius (large-R) jet that differs from a large-R jets originating from light-quarks or gluon for a distinct radiation pattern. 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 order to suppress contribution from pile-up, the large-R jets undergo a *trimming* procedure [92] 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 for the jets collection adopted in the analysis described in this thesis 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 large-R jet mass is the result of a combination of calorimeter and tracking information [57]. In fact, the mass of a jet is typically defined as the invariant mass of the jet constituents, which are assumed to be massless:

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

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

$$m^{\text{TA}} = \frac{p_T^{\text{calo}}}{p_T^{\text{track}}} \times m^{\text{track}} \tag{3.8}$$

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

$$m^{\text{comb}} = w_{\text{calo}} m^{\text{calo}} + w_{\text{TA}} m^{\text{TA}} \tag{3.9}$$

is exploited as jet mass definition which suits all p_T regimes [54].
In the analysis described in this thesis, the large-R jets considered are formed using Track-CaloCluster (TCC) [60] algorithm. The idea of TCC is to maximally exploit the strengths of tracking detectors and calorimeters for improved measurements of hadronic interactions in highly energetic hard-scatter processes.
The ATLAS calorimeter has an excellent energy resolution, which improves as the energy is increased until reaching a constant value, so long as the full hadronic shower is contained within the calorimeter. However, the granularity of the calorimeter is insufficient to resolve the angular separation between highly boosted hadronic decays of massive particles, and thus the angular resolution can be a limiting factor. In contrast, the ATLAS tracking detector has an excellent angular resolution, but as the tracks become more energetic, they are less curved by the magnetic field and the transverse momentum resolutions degrades. This forms the basis of the TCC. In order to maximally benefit at high pT from the superb calorimeter energy resolution and excellent tracker spatial resolution, a 4-vector is formed which (to first order) uses topo-clusters for the scale components (pT, m - where the mass of a single topo-cluster in ATLAS is defined to be zero) while using the track parameters to determine the angular (η, ϕ) coordinates.

Reconstruction

In order to build 4-vectors from a combination of track and topo-cluster information, first a track-cluster matching criterion is defined. The algorithm attempts to match every good quality track to every topo-cluster following two steps. In the first stage, the uncertainty on the track extrapolation to the calorimeter is compared to the width of the topo-cluster and if the extrapolation uncertainty is larger than the topo-cluster width, then the track is discarded from the matching procedure. Otherwise, the matching continues to the second step, where a track-cluster pair is defined as matched whenever their angular separation $\Delta R < \sqrt{\sigma^2_{track} + \sigma^2_{cluster}}$, with $\sigma_{cluster}$ the topo-cluster width and σ_{track} the track extrapolation uncertainty. Once tracks and topo-clusters have been matched, the actual TCC four-vectors are built. In Figure 3.10 a schematic example is shown, in order to assist in understanding the procedure. When referring to the figure, (1) refers to TCC object 1, c_1 refers to topo-cluster c_1 and t_1 refers to track t_1. For a direct single match between a track from the selected hard scatter vertex and a topo-cluster, such as (1) in the figure, the topo-cluster energy and the track direction are used to form a single TCC:

$$TCC_{(1)} = (p_T^{c_1}, \eta^{t_1}, \phi^{t_1}, m^{c_1} = 0) \tag{3.10}$$

In the case of topo-clusters not matching any tracks, such as (2), the topo-cluster four-vector is directly used to create a TCC:

$$TCC_{(2)} = (p_T^{c_7}, \eta^{c_7}, \phi^{c_7}, m^{c_7} = 0) \tag{3.11}$$

Independent tracks from the selected hard scatter vertex which do not match any topo-clusters, such as (3) in the figure, are also used directly to create TCCs:

$$TCC_{(3)} = (p_T^{t_6}, \eta^{t_6}, \phi^{t_6}, m^{t_6} = 0) \tag{3.12}$$

Once there is a match between multiple tracks and a single cluster, multiple topo-clusters and a single track, or especially multiple topo-clusters with multiple tracks, as in the case of (6) and (7), the situation becomes more complex. The TCC reconstruction procedure always creates exactly one TCC object per track originating from the primary vertex, where the track angular coordinates are used, but the scale coordinates must be adapted to account for energy sharing between the different matches. For each track used to seed a TCC, all matching topo-clusters are found. For each of those matching tracks, the energy is then divided between all of the tracks which match that cluster, with the split defined by the fraction of p_T contributed by a given track compared to all of the other matching tracks. As a simple example where there are two tracks matching a single topo-cluster, let's consider the case of (4) and (5). The TCC four-vectors in these cases are:

$$TCC_{(4)} = (p_T^{c_2} \frac{p_T^{t_2}}{p_T(p^{t_2} + p^{t_3})}, \eta^{t_2}, \phi^{t_2}, m^{c_2} \frac{p_T^{t_2}}{p_T(p^{t_2} + p^{t_3})} = 0) \tag{3.13}$$

$$TCC_{(5)} = (p_T^{c_2} \frac{p_T^{t_3}}{p_T(p^{t_2} + p^{t_3})}, \eta^{t_2}, \phi^{t_2}, m^{c_2} \frac{p_T^{t_3}}{p_T(p^{t_2} + p^{t_3})} = 0) \tag{3.14}$$

Whether the jet is built from topo-clusters, all TCCs, or only combined TCCs, a calibration using dedicated correction factors is needed to account for the non-uniform and non-compensating energy response of the ATLAS calorimeters. Topo-cluster jets are fully calibrated using both Monte Carlo derived jet energy (JES) and mass scale correction (JMS).

The different stages of the large-R jet calibration procedure are the following. 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 [57]. Finally, the jets are calibrated *in situ* (Figure 3.11) using response measurements in data. The calibration is done taken into account the response of the detector in the measure of the energy [59] (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.

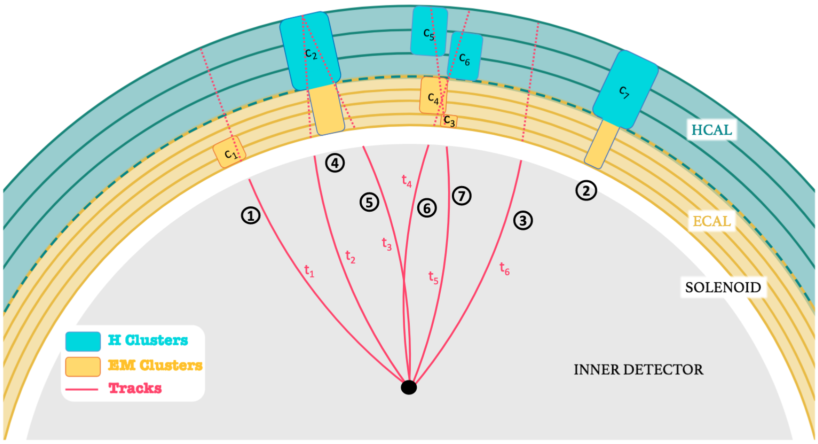


Figure 3.10: A schematic view of seven TCC objects representing (1) a simple track-cluster match, (2) a topo-cluster without a matching track, (3) a track without a matching cluster, (4) and (5) are each tracks matching a single cluster but sharing that cluster's energy, and (6) and (7) showing a much more complex scenario with multiple track-cluster matches. [60]

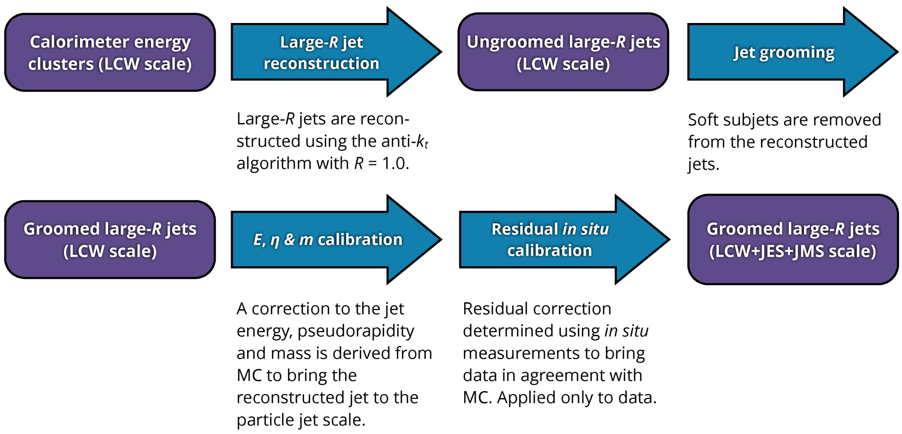
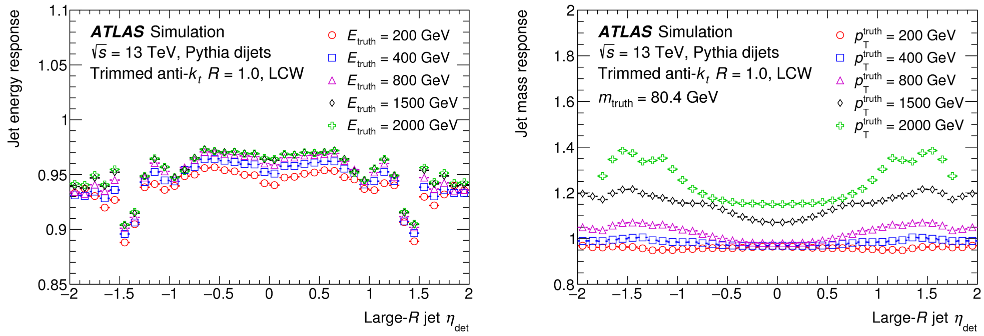


Figure 3.11: Overview of the large-R jet reconstruction and calibration procedure. The calorimeter energy clusters from which jets are reconstructed have already been adjusted to point at the event primary hard-scatter vertex.

3.4.4 Tagging b-jets

An important role is played by the jet *flavor tagging* task in the identification of the jets. The identification of jets containing b-hadrons (*b-jets*) against the large jet background containing c-hadrons (*c-jets*) or containing neither b- or c-hadrons



(a) (b)

Figure 3.12: Jet energy response (a) and jet mass response (b) as a function of η of the jet. [11]

(*light-flavour jets*) is of major importance in many areas of the physics programme of the ATLAS experiment. It has been decisive in the observations of the Higgs boson decay into bottom quarks [10] and of its production in association with a top-quark pair [9] and plays a crucial role in a large number of Standard Model precision measurements, studies of the Higgs boson properties and searches for new phenomena.
The ATLAS Collaboration uses various algorithms to identify b-jets[13] when analysing data recorded during Run 2 of the LHC (2015–2018). These algorithms exploit the long lifetime, high mass and high decay multiplicity of b-hadrons as well as the properties of the b-quark fragmentation. Given a lifetime of the order of $1.5\,\mathrm{ps}$ ($< c\tau >\approx 450\,\mu\mathrm{m}$), measurable b-hadrons have a significant mean flight length $< l >= \beta\gamma c\tau$ in the detector before decaying, generally leading to at least one vertex displaced from the hard-scatter collision point.
The strategy for b-tagging follows a two-stage approach. Firstly, low-level algorithms reconstruct the characteristic features of the b-jets. Secondly, in order to maximise the b-tagging performance, the results of the low-level b-tagging algorithms are combined in high-level algorithms consisting of multivariate classifiers. The performance of a b-tagging algorithm is characterised by the probability of tagging a b-jet (b-jet tagging efficiency, ϵ_b) and the probability of mistakenly identifying a c-jet or a light-flavour jet as a b-jet, labelled ϵ_c or ϵ_l, or equivalently in terms of c-jet and light-flavour jet rejections, defined as $1/\epsilon_c$ and $1/\epsilon_l$, respectively. Mostly two algorithms have been adopted in the Run 2 data analyses: *MV2* and *DL1* [74]. The former is a boosted decision tree (BDT) algorithm that combines the outputs of low-level tagging algorithms. It is trained on the hybrid $t\bar{t}\ +\ Z'$ sample and, to avoid differences in the kinematic distributions of signal (b-jets) and background (c-jets and light-flavour jets), the b-jets and c-jets are reweighted

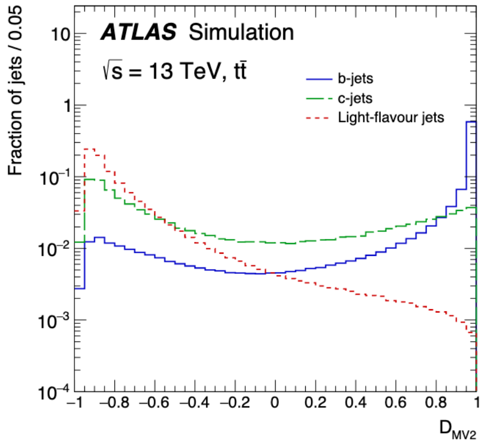
in p_T and $|\eta|$ to match the spectrum of the light-flavour jets. For training, the c-jet fraction in the background sample is set to 7%, with the remaining composed of light-flavour jets, in order to allow the charm rejection to be enhanced whilst preserving a high light-flavour jet rejection. The output discriminant of the MV2 algorithm for b-jets, c-jets and light-flavour jets evaluated with the baseline $t\bar{t}$ simulated events are shown in Figure 3.13a.
The second mentioned b-tagging algorithm, DL1, is based on a deep feed-forward neural network (NN). The multidimensional output correspond to the probabilities for a jet to be a b-jet, a c-jet or a light-flavour jet. The input variables to DL1 consist of those used for the MV2 algorithm with the addition of the some c-tagging variables related to the dedicated properties of the secondary and tertiary vertices (distance to the primary vertex, invariant mass and number of tracks, energy fraction and rapidity of the tracks associated with the secondary and tertiary vertices). A jet p_T and $|\eta|$ reweighting similar to the one used for MV2 is performed.
During training all flavours are treated equally, therefore the network can be used for both b-jet and c-jet tagging. The DL1 b-tagging discriminant is defined as:

$$D_{DL1} = ln(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_l}) \tag{3.15}$$

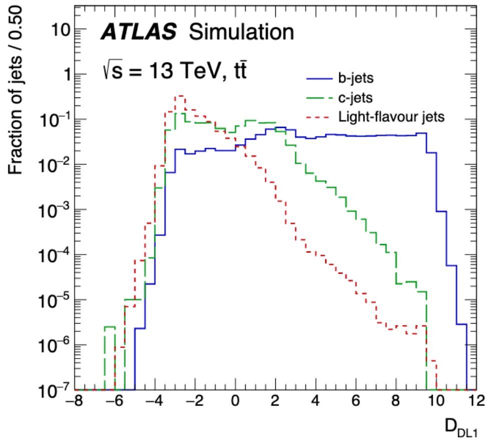
where p_b, p_c , p_l and f_c represent respectively the b-jet, c-jet and light-flavour jet probabilities, and the effective c-jet fraction in the background training sample. Using this approach, the c-jet fraction in the background can be chosen a posteriori in order to optimise the performance of the algorithm. The output discriminant of the DL1 algorithm for b-jets, c-jets and light-flavour jets evaluated with the baseline $t\bar{t}$ simulated events are shown in Figure 3.13b.
B-tagging on large-R jets is applied on VR track-jets *ghost associated* [42] with the large-R jet. There are several benefits of using track jets for finding b-hadrons. As b-tagging pattern recognition algorithms only use information from the inner detector, they can be decoupled from calorimeter jet finding in order to identify b-hadrons. The track jet algorithm can be optimized and calibrated specifically for b-tagging, independent of the calorimeter jet algorithm, which can be optimized separately to interpret the hadronic final state. Using the ghost-association technique, the b-tagging information provided by the track jets can be integrated with the calorimeter jet algorithm whilst taking full advantage of the calorimeter jet (and subjet) structure and shape.

The identification of boosted Higgs bosons decaying in two b-jets

The identification of boosted Higgs boson decaying via the dominant $H \to b\bar{b}$ mode is of primary important at LHC.



(a)



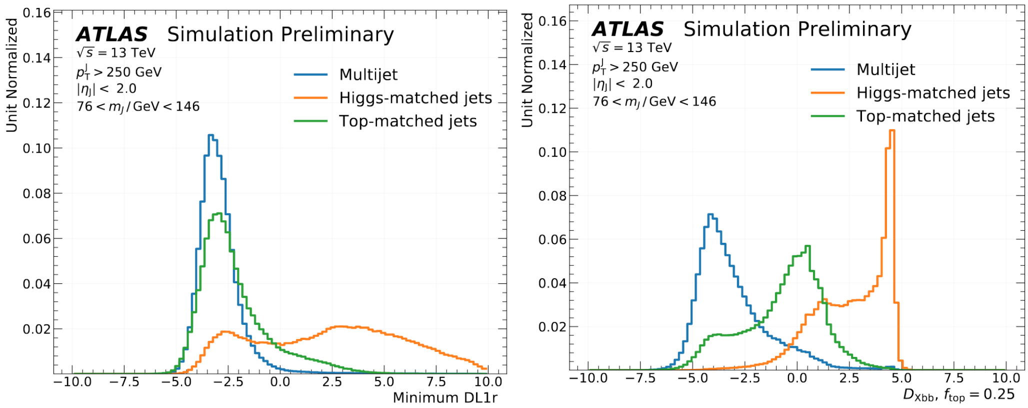
(b)

Figure 3.13: Distribution of the b-tagging (a) MV2 and (b) DL1 outputs for b-jets, c-jet and light-flavour jets in the baseline $t\bar{t}$ simulated events [10].

The baseline double b-tagging method requires two track-jets within a large-R jet, each with a b-tagging discriminant over some threshold, inevitably leading to a lower tagging efficiency. To take full advantage of the information in the large-R jet, a neural network algorithm which combines the outputs from the high-level single-b tagging algorithms with large-R jet kinematics has been developed [62]. Flavor tagging information is supplied by the DL1r algorithm (a variant of DL1 introduced above, trained on VR track-jets), which outputs three values corresponding to the probabilities for the jet to be a b-, c-, or light-jet. Discriminants from up to three leading VR track-jets are passed to the double b-tagging algorithm. In addition to this b-tagging information, the p_T and η of the large-R jets are used, for up to eleven input variables for each jet. For training, distributions of the transverse momentum of the large-R jets are downsampled, in that jets are removed until the shape of the p_T distribution matches for all processes, preventing the model from learning the different kinematic properties of the jets.
The type of network adopted for this task is a multi-class feed-forward neural network with three outputs: respectively the probabilities for the large-R jet to be produced by an Higgs decay (p_{Higgs}), a multijet event ($p_{multijet}$) or a top decay (p_{top}). These probabilities can be combined into a single discriminant roughly corresponding to a log-likelihood ratio, defined as follows:

$$D_{H_{bb}} = ln \frac{p_{Higgs}}{f_{top} \cdot p_{top} + (1 - f_{top}) \cdot p_{multijet}} \tag{3.16}$$

where f_{top} determines the fraction of top background and can be determined a posteriori, after the training, according to the physics case where it is used.
Figure 3.14 shows the $D_{H_{bb}}$ discriminant calculated for jets in the Higgs, top and multijet samples. For comparison, the discrimination obtained by directly using the DL1r outputs of the subjets is shown as well. The double b-tag requirement is fulfilled when the minimum DL1r discriminant value for the two highest p_T subjets is above a given threshold.

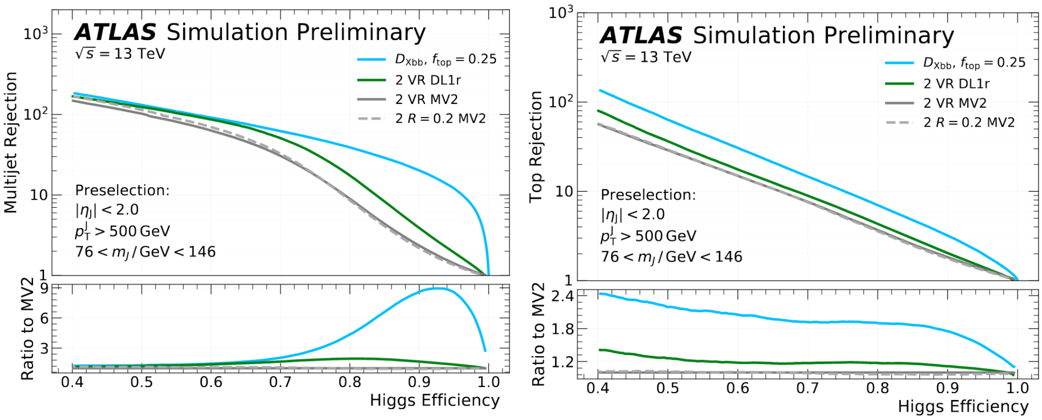


(a) (b)

Figure 3.14: The discriminant distributions (normalized to unity), for the DL1r-based benchmark defined as the minimum of the two leading VR track-jets discriminant (a), and the double b-tagging algorithm $D_{H_{bb}}$ with a top fraction of f=0.25 (b) [62].

Figure 3.15 shows the signal efficiency and corresponding background rejection for a wide range of possible values of this threshold, evaluated for either multijet or top jet backgrounds. Jets are considered tagged when either $D_{H_{bb}}$ or the minimum of the two VR track-jets discriminant DL1r (MV2 is shown as well) is above some fixed threshold. f_{top} here is set to 0.25.
When compared to double b-tagging algorithms used by ATLAS in Run 2, $D_{H_{bb}}$ improves discrimination against top and multijet backgrounds. Relative to a double MV2 tag, the multijet (top) rejection at 60% efficiency is increased by a factor of 1.4 (2.0) for jets with $p_T > 500\,\mathrm{GeV}$. The improvement is still substantial when comparing to a double DL1r tag: the multijet rejection is roughly equal, while top rejection is improved by a factor of 1.7. Furthermore, the $D_{H_{bb}}$ discriminant offers these performance increases over DL1r without significantly worsening the impact on the shape of the jet mass distribution. Additionally, track jets are explicitly chosen to originate from the primary vertex, significantly reducing the performance dependence on pileup. This is especially important when reconstructing low p_T

b-hadrons that can be present in highly boosted states if the b-hadron is produced in the opposite direction to the boost of the decaying particle. These relatively low-p_T b-hadrons might otherwise have been lost due to a higher p_T-threshold imposed on calorimeter jets in order to reduce pileup.



(a) (b)

Figure 3.15: Multijet (a) and top jet (b) rejection as a function of the $H \rightarrow b\bar{b}$ tagging efficiency, for large-R jet $p_T > 500\,\mathrm{GeV}$. Performance of the $D_{H_{bb}}$ algorithm is compared to DL1r and to two variants of MV2, one evaluated on variable-radius (VR, $\rho = 30\,\mathrm{GeV}$) jets, the other on fixed-radius (FR, R = 0.2) jets. The efficiency and rejection are calculated with respect to jets that passed some p_T, η, and mass preselection requirements [62].

3.4.5 Boson tagging

The capability to understand and discriminate the large-R jet substructure is crucial for a large variety of physics analyses with boosted bosons, where the reconstruction efficiency of the hadronic boson decays is high if using large-R jets instead of two separate small-R jets. Boson tagging has the aim to select large-R jets consistent with vector boson hadronic decays $V \rightarrow qq$ $(V = Z/W)$.
A function exploiting the N-prongness of the jet is the energy correlation function (ECF), defined as:

$$ECF(N, \beta) = \sum_{i_1, i_2, \ldots, i_N \in J} \left(\prod_{a=1}^{N} p_{Ti_a}\right)\left(\prod_{b=1}^{N-1} \prod_{c=b+1}^{N} \theta_{i_b i_c}\right)^{\beta} \tag{3.17}$$

Here, the sum runs over all particles within the system J (either a jet or the whole event). Each term consists of N energies multiplied together with $\binom{N}{2}$ pairwise angles raised to the angular exponent β. 3.18 is most appropriate for e+e colliders where energies and angles are the usual experimental observables. For hadron colliders, it is more natural to define the ECF as a transverse momentum correlation function:

$$ECF(N,\beta) = \sum_{i_1,i_2,\ldots,i_N \in J} (\prod_{a=1}^{N} p_{Ti_a})(\prod_{b=1}^{N-1} \prod_{c=b+1}^{N} \Delta R_{i_b i_c})^{\beta} \tag{3.18}$$

where $\Delta R_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. From this, the ratio between two energy correlation functions, which is dimensionless, can be defined as follows:

$$D_N^{\beta} = \frac{ECF(N+1,\beta)(ECF(N-1,\beta)}{ECF(N,\beta)^2} \tag{3.19}$$

The energy correlation double ratio effectively measures higher-order radiation from leading order (LO) substructure. For a system with N subjets, the LO substructure consists of N hard prongs, so if D_N is small, then the higher-order radiation must be soft or collinear with respect to the LO structure. If D_N is large, then the higher-order radiation is not strongly-ordered with respect to the LO structure, so the system has more than N subjets. Thus, if D_N is small and D_{N1} is large, then we can say that a system has N subjets.

In the case of 2 prongness, such as for the boson decays in two quarks the function adopted to distinguish between jets originating from a single parton and those coming from the two-body decay of a heavy particle is D_2, optimized with $\beta = 1$. In particular, lower D_2 values are indicative of two-prong large-R jets. More information are given in [94].
This variable has often been used to tag the boson decay in two quarks and, along with the jet mass and the number of tracks associated to the large-R jet, it is used to define the boson tagger [61] adopted by many Run-2 physics analyses.

CHAPTER 4

Full-hadronic YXH analysis

In this chapter the actual analysis work carried out for this Thesis is described. It is organized as follows: first the analysis is introduced, then the event selection is described followed by the background estimation and the signal mass resolution studies; after these, the systematic uncertainties are treated and finally the statistical treatment is discussed. Results are shown in the last section.

4.1 Introduction

The sensitivity of the Higgs mass to radiative corrections implies either extreme fine-tuning in the model or the existence of new physics at an energy scale not far above the Higgs boson mass. This theoretical motivation, coupled with the existing experimental mass reach of the Large Hadron Collider (LHC) at CERN, motivates searches for new particles at the TeV scale. Because the Higgs boson couples to mass, it is natural to expect that these new heavy particles may have decays to a Higgs boson.

A search for a new TeV-scale narrow-width boson Y, which decays to a Standard Model Higgs boson H and a new particle X with a mass on the weak scale, where a fully hadronic final state is assumed for both particles, has been carried out, is described in this Thesis and is documented in [7].

The analysis is sensitive to X masses spanning several orders of magnitude, from $\mathcal{O}(10)$ GeV to $\mathcal{O}(1)$ TeV.

Figure 4.1 shows the Feynman diagram for this process, where the X can have a variety of hadronic decays. The masses of the parent and daughter particles yield a kinematic scenario where the final state particles are highly Lorentz-boosted, motivating a reconstruction using large-R jets and the use of jet substructure to

distinguish the boson decay products.
The main highlight of this analysis is the introduction of an unsupervised machine learning architecture used for *anomaly detection* on jets to select *X* particles based solely on their incompatibility with the expected SM background. It is the first time ATLAS publishes a work with a fully unsupervised method.
A modified Heavy Vector Triplet (Section 1.3.1) framework is adopted only as a benchmark model to set upper limits on production cross section. For this latter target, a reconstruction of the X particle with two small-R jets is used to recover sensitivity to a topology where X is less boosted, significantly extending the region of sensitive phase space.

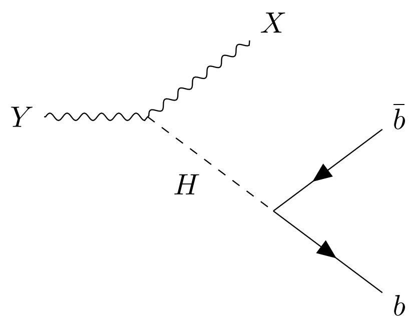


Figure 4.1: Feynman diagram of the considered signal process. The Y is produced in the initial pp collision and decays to a fully hadronic final state via a SM Higgs boson $H \to b\bar{b}$ and a new particle X. The only assumption on the X decay is that it decays hadronically. [7]

The search exploits the full LHC Run 2 $\sqrt{s} = 13$ TeV pp dataset, collected by the ATLAS detector from 2015 to 2018 and corresponding to an integrated luminosity of $139\,\text{fb}^{-1}$. A previous search was performed by ATLAS using $36\,\text{fb}^{-1}$ under the assumption of $X \to q\bar{q}$, with no significant excess found covering Y masses from 1 to 4 TeV and X masses from 50 to 1000 GeV [32]. In addition to the increased luminosity of the dataset, this result includes several key improvements with respect to this last iteration, such as a neural net-based tagger optimized for the boosted H_{bb} topology, anomaly detection for enhanced signal model independence, and the usage of two orthogonal regions to capture both boosted and resolved reconstruction of the nominal X decay to two quarks.

This thesis particularly focuses on the background estimation because its development represents most of the work carried out over the three years of PhD.

4.2 Data and Simulated Samples

The analysis has exploited the full run-2 dataset. In this section information regarding Data sample and Monte Carlo simulations used for signal interpretations is given.

4.2.1 Data

The analyzed dataset used corresponds to 139 fb^{-1} of LHC pp collision data collected by the ATLAS detector from 2015 to 2018.
The exact integrated luminosity, separated according to data taking periods, correspond to 3.2 fb^{-1} from 2015, 33.0 fb^{-1} from 2016, 44.3 fb^{-1} from 2017, and 58.5 fb^{-1} from 2018. Data have been selected from the the good runs lists (GRL), ensuring that all the relevant elements of the ATLAS detector were fully operational and efficient while the data were collected.

Events are further selected with a single-jet trigger, where events are required to have a jet at trigger-level with a p_T that exceeds a certain value, depending on the data period. The trigger requests are listed here:

- 2015: `HLT_j360_a10_lcw_sub_L1J100`, $p_T \geq 360$ GeV
- 2016: `HLT_j420_a10_lcw_L1J100`, $p_T \geq 420$ GeV
- 2017 & 2018: `HLT_j460_a10t_lcw_jes_L1J100`, $p_T \geq 460$ GeV

Selections are imposed (Section 4.3.1) that ensure these triggers are only used after reaching an efficiency plateau, avoiding the effect of the trigger turn-on in the analysis.

4.2.2 Monte Carlo simulation

Simulated events are generated with a variety of Monte Carlo (MC) generator processes that run in stages. The pp hard scatter physics process is simulated, and the final state particles are subsequently showered and decayed. This full description of the event is then propagated through a detailed detector simulation based on the software GEANT4 [19].
The MC simulation is weighted to match the distribution of the average number of interactions per bunch crossing μ observed in collision data. All simulated samples included in this analysis were produced with three different campaigns: `mc16a` corresponds to 2015-2016 data-taking conditions, `mc16d` to 2017, and `mc16e` to 2018. These three campaigns are weighted to the integrated luminosities of their

respective data-taking periods and combined to produce simulation for the entire Run 2 dataset. Simulated events are reconstructed with the same algorithms run on collision data.

Signal Samples

The adopted baseline for the $Y \to XH$ signal samples is the Heavy Vector Triplet (HVT) Model A (Sec. 1.3.1), for which the branching fractions to fermions and gauge bosons are comparable. The standard WH configuration is modified to replace the W with the X particle. The X is considered to be charged with variable mass, natural width of 2 GeV (smaller than the detector resolution), and spin-1, with allowed decays only to $u\bar{d}$ with 100% branching ratio. Samples with resonance Y masses between 1-6 TeV are generated using MadGraph5 [25] interfaced to Pythia 8.244 p3 [58] for shower and hadronization with NNPDF23LO PDF [41] and the ATLAS A14 [30] tune to underlying-event data.
The simulated width of the Y ranges between 10s to 100s GeV, calculated in narrow-width approximation and at tree-level.

All signals are generated with an arbitrary cross section of 1 pb^{-1}. The generated mass points are illustrated in Figure 4.2.

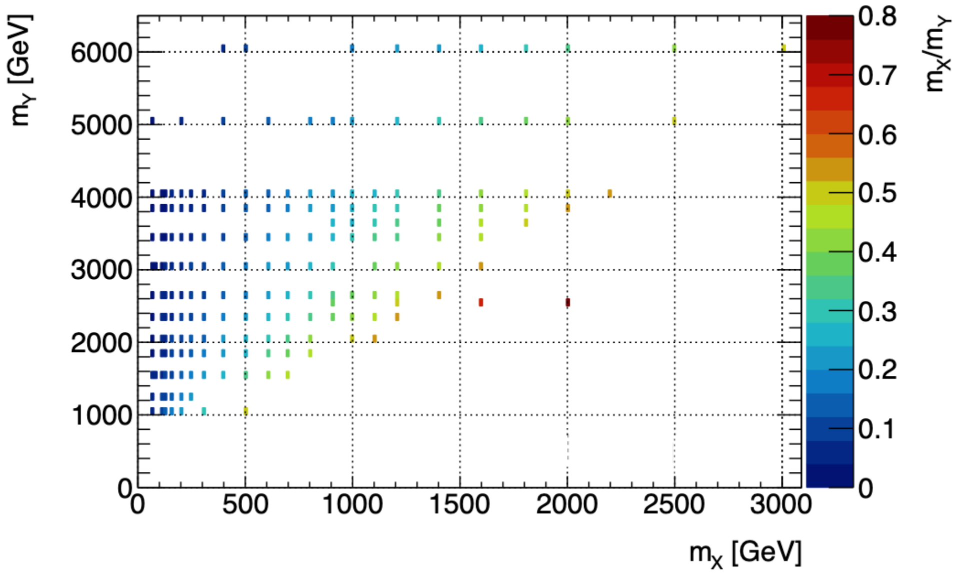


Figure 4.2: Diagram of simulated signal points used in the analysis, each defined by a value of m_X and m_Y.

Some additional R&D signals are used in the analysis to assess the model-independence of the anomaly region (Section 4.3.3). Heavy-flavor and three-prong X candidates are studied via Pythia-8 generated processes[1]. Both signals have a 3000 GeV parent particle decaying to daughters of masses 200 and 400 GeV, ensuring large-R jet reconstruction. For the heavy-flavor decay, the process is $A^0 \rightarrow$ HZ, where the Z and H are non-SM bosons that both decay to $b\bar{b}$. To produce a three-prong jet, the process is W' $\rightarrow$ WZ, where the W and Z are non-SM bosons that both decay to three light quarks.
The third signal is a dark jet of mass 3500 GeV, produced non-resonantly with the Pythia Hidden Valley module using ModelA.

Background Samples

Though no background simulation is formally used in the analysis, some background MC is studied for analysis optimization.
Dijet processes are simulated with PYTHIA8 using the NNPDF23LO PDF set and the A14 tune. The samples are generated in approximate slices in bins of momentum transfer, to ensure high statistics across the momentum spectrum.
Top backgrounds are simulated using POWHEG [24], and as with signal are interfaced to PYTHIA8 with NNPDF23LO PDF and the A14 ATLAS tune. Samples are produced in four slices of the di-top invariant mass to increase statistics in the lower cross section areas of phase space.
Single bosons (W/Z) produced in association with jets are modeled with SHERPA V2.1.1 [81] with the CT10 generator tune/PDF. Only hadronic decays of the W and Z are included. Similar to the top processes, samples are simulated in three slices of the boson p_T to ensure sufficiently well-populated tails of phase space.

4.3 Objects identification and Event selection

The experimental signature of the $Y \rightarrow XH$ signal contains at least two jets with high transverse momentum. Let's first define the physics objects (introduced in Chapter 3) and then let's move to the description of the event selection.

The large-R jets used in the analysis are taken from Track-CaloCluster jets collection. In order to further reduce the pileup effect, a grooming technique is applied to fatjet with the trimming algorithm with subjet radius 0.2 and fraction of the p_T cut of 0.05. Other requirements applied are: 1) jet $|\eta| < 2.0$, to ensures a good overlap with the tracking volume of ATLAS detector; 2) leading jet $p_T >$ 200 GeV, to assure resonance p_T is large enough to collect its hadronic decays in the boosted

[1]Used for R&D in https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnomalyBumpHuntRound2

cone jet.

Variable-radius track jets are used to identify b-quark-induced jets in the merged analysis. They are built by clustering Inner Detector tracks using the anti $-$ k_T algorithm. Tracks with p_T greater than 500 MeV and passing a loose set of cuts are required to be associated with the primary vertex of the event, defined as the vertex with the largest $\sum p_T^2$. The ρ parameter is set to 30 GeV and the minimum and maximum values for the radius are set to 0.02 and 0.4, outside of which the jet's radius stays fixed to avoid jets with an arbitrary small or large radius parameter.

Small radius jets are based on particle flow jet constituents and are used in the resolved regime, where X resonance can be reconstructed from two separated jets with a fixed radius parameter of 0.4, $|\eta| < 2.5$ and $p_T >$ 20 GeV.

Leptons are not explicitly used in the analysis, but they are here discussed since they take part in the overlap removal, described below. Electrons are required to have $p_T > 7$ GeV and $|\eta| < 2.47$. Muons are required to have $p_T > 7$ GeV and $|\eta| < 2.7$. To ensure that leptons originate from the interaction point, transverse and longitudinal track impact parameter criteria, with respect to the beam line, are imposed, $|d_0/\sigma(d_0)| < 3(5)$ for muons (electrons) and $|z_0 \sin\theta| < 0.5$ mm respectively. No lepton isolation criteria are applied.

The reconstruction of the same energy deposits as multiple objects is resolved. The procedure is applied to the selected leptons and jets. If two electrons share the same track, or the separation between their two energy clusters satisfies $|\Delta\eta| < 0.075$ and $|\Delta\phi| < 0.125$, then the lower-p_T electron is discarded. Electrons that fall within $\Delta R = 0.02$ of a selected muon are also discarded. For nearby electrons and small-R jets, the jet is removed if the separation between the electron and jet satisfies $\Delta R < 0.2$; the electron is removed if the separation satisfies $0.2 < \Delta R < 0.4$. For nearby muons and small-R jets, the jet is removed if the separation between the muon and jet satisfies $\Delta R < 0.2$ and if the jet has less than three tracks or the energy and momentum differences between the muon and the jet are small; otherwise the muon is removed if the separation satisfies $\Delta R < 0.4$. To prevent double-counting of energy from an electron inside the large-R jet, the large-R jet is removed if the separation between the electron and the large-R jet satisfies $\Delta R < 1.0$. All conditions are summarized in Table 4.1.

Table 4.1: Conditions of the overlap removal. These cuts are applied sequentially.

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

4.3.1 Preselection

Before a selection based on physics reasons, a veto is applied to reject events where there is an error in any detector system. To reject non-collision backgrounds, such as calorimeter noise, beam halo interactions, or cosmic rays, the standard ATLAS event cleaning procedure is applied: events are required to have at least one primary vertex, selected as the vertex with the highest sum of track p_T^2 associated to it that also has a minimum of two associated tracks.

After this, only events with at least two large-R jets are selected. The two p_T-leading large-R jets in the event are considered as Higgs and X boson candidates. Then a preselection is applied to isolate the phase space of events that will be used for both the signal region and the estimation of background in that region, in order to significantly skim the dataset used for the real analysis. It consists in the following cuts:

- Either leading or sub-leading jet mass > 50 GeV
- Leading large-R jet $p_T > 500$ GeV
- Invariant mass of the two leading large-R jets m_{JJ} (= m_{XH}) > 1.3 TeV

The last selection has been applied to avoid trigger turn-on effects or, in other words, to be in the plateau of the trigger selection efficiency. The effectiveness of this cut has been demonstrated in another analysis with the fully hadronic final state (resonant VV $\rightarrow$ JJ [35], with V=W,Z) and inherited from there.

4.3.2 X/H candidate identification

After preselection cuts, an ambiguity resolution is required to determine which of the two p_T-leading large-R jets in the event is more likely to be the Higgs boson and thus subject to the Higgs boson selection criteria. Once this ambiguity is resolved, the selection on the jets diverges: $D_{H_{bb}}$-tagging is applied to the H candidate and X-tagging is applied to the X candidate (Section 4.3.3). The H and X candidates are selected only between the two p_T-leading large-R jets, any eventual additional jet is not considered.
This selection is based on the neural net-based $D_{H_{bb}}$classifier (defined in Section 3.4.4), with $f_{top} = 0.25$, which separates bosons decaying to $b\bar{b}$ from top quark and QCD jets. The version used in the analysis includes a reweighting of all training inputs to have the same p_T and η distributions, to minimize bias of the tagger to high-p_T or central jets respectively.
To perform the large-R jet ambiguity resolution, $D_{H_{bb}}$ is computed for both the p_T-leading large-R jets in the event.
To ensure optimal and accurate performance of the ambiguity resolution, several potential procedures were studied. They were compared in terms of their accuracy in selecting the Higgs candidate when compared to truth Higgs association, where the accuracy is defined as the fraction of events where the jet selected by the ambiguity resolution algorithm is the same as the jet that is found to be ΔR-matched to the truth Higgs hadron.
The ambiguity resolution algorithm selected for this analysis is referred to as **Scheme A**: it selects as the Higgs candidate the large-R jet (of the two p_T-leading in the event) with the highest $D_{H_{bb}}$output score.
Scheme A is compared to four other schemes:

- **Scheme B**: as a benchmark for comparison, the assignment procedure from the early 36.1 fb^{-1}paper is replicated, which used in order of priority the jet mass, number of DL1r small-R b-tags, and jet p_T to determine the H candidate. Specifically, if exactly one of the leading jets is in the Higgs mass window of 75 to 145 GeV, it is the H candidate. If neither or both jets are in the Higgs mass window, the jet with the highest number of DL1r b-tags is the H candidate. If both jets have the same number of b-tags, the higher p_T jet is the H candidate.

- **Scheme C**: to first use the large R jet mass to make the X/H determination, but replace the reliance of the early paper procedure on DL1r with the $D_{H_{bb}}$. If there is a jet in the event with a mass in the Higgs mass window, that jet is the H candidate. If neither or both of the leading jets is in the Higgs mass window, the leading $D_{H_{bb}}$jet is the Higgs candidate.

- **Scheme D**: similar to the early paper Scheme B, but replacing the number of DL1r (Section 3.4.4) small-R b-tags with the number of fixed efficiency $D_{H_{bb}}$tags (a calibrated quantity). If both/neither of the jets is in the Higgs mass window and the number of $D_{H_{bb}}$-tags is equal, the Higgs candidate is selected as the lower p_T jet.
- **Scheme E**: the final reliance of Scheme D on the jet p_T was found to introduce unwanted kinematic biases in the background estimation. Scheme E is identical to Scheme D, except at the last resolution stage, the Higgs candidate is assigned randomly, rather than as the lower p_T jet.

Figure 4.3 demonstrates the accuracy of this procedure for correctly selecting the Higgs jet from the two leading jets as a function of m_X/m_Yin the signal grid. As

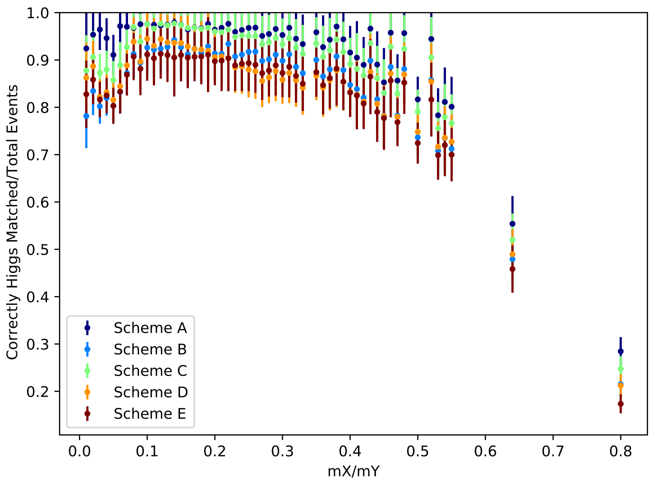


Figure 4.3: Percent events where the selected Higgs candidate is matched to a true Higgs jet for several candidate procedures. The one selected in this analysis is Scheme A.

shown in Figure 4.3 the accuracy decreases for $mX/mY > 0.5$ and it is due to the fact that Higgs truth jet for these signals is in most of the cases the third or even subsequent ones. For this reason this analysis investigates signals up to mass ratio m_X/m_Y= 0.5.

As Scheme A is found to be the most accurate as determining the Higgs jet from a choice of the leading two, it is chosen and applied as the first step in the analysis after preselection.
Scheme A relies on the use of the uncalibrated $D_{H_{bb}}$ score to distinguish between jets, therefore a potential discrepancy between data and MC may be introduced.

The only way for this discrepancy to affect the analysis is if both jets pass the $D_{H_{bb}}$tagging process (both are good H candidates and the two jets may have been switched in the ambiguity resolution procedure): in this case, the impact of $D_{H_{bb}}$tagging efficiency discrepancies could affect the final efficiency for signal and data. To assess the potential size of such an impact, the fraction of signal events where both X and H candidates pass the 60% $D_{H_{bb}}$ tagging working point has been checked. In Figure 4.4 the fraction of events with both X and H jets passing the 60% WP on $D_{H_{bb}}$is shown, for all YXH signals available in our search grid. It can be seen that <0.5% of the events satisfies this request, leading to a negligible impact on the efficiency loss due to a wrong assignment. This study suggests that our assignment criterion is stable against the uncalibration problem, without any significant loss on signal efficiency.

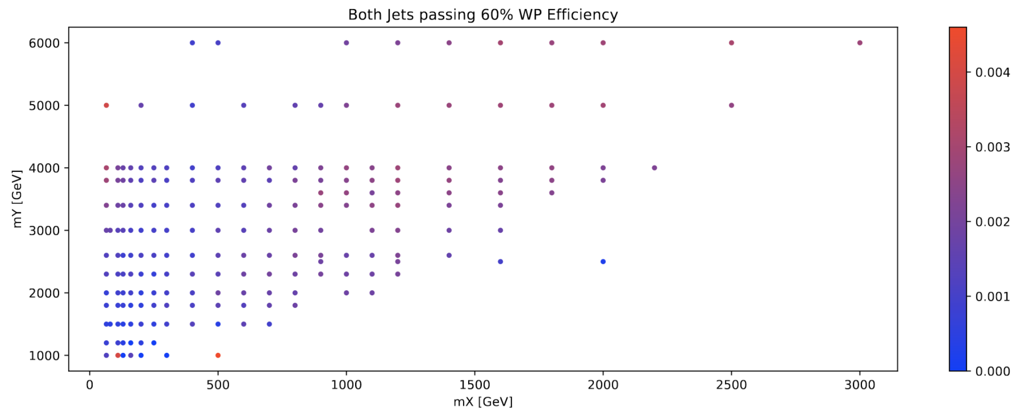


Figure 4.4: Fraction of events with both leading large-R jets passing the 60% WP on $D_{H_{bb}}$.

The resulting distribution of $D_{H_{bb}}$ for the large-R jet chosen as the Higgs boson candidate, for both data and representative $Y \to XH$ signals at preselection, is shown in Figure 4.5. Additionally, a preselection of $D_{H_{bb}} > -2$ is applied to remove events that are determined to be not $H \to b\bar{b}$ -like, thus ensuring the data-driven background estimation focuses on an area of phase space that is close to that of the signal.

4.3.3 X candidate selection

Once the Higgs candidate is selected, the other of the two p_T-leading large-R jets automatically becomes the X candidate.
A selection is placed on the X jet to provide additional rejection of multi-jet and

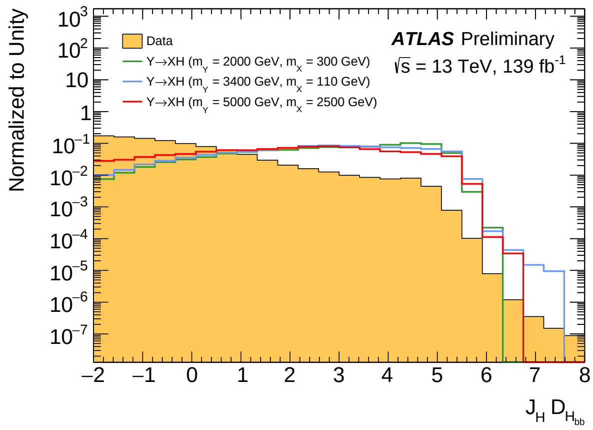


Figure 4.5: Distributions of the H candidate $D_{H_{bb}}$ score in data after preselection requirements are applied. Also shown are three $Y \to XH$ simulated signals, labelled by the masses of the Y and X particles. All distributions are normalized to unity. [7]

other backgrounds. Three signal regions are defined in this analysis, each using a different method of distinguishing the X based on jet substructure. The first two are used to give exclusion results on the specific $Y \to XH$ signal hypotheses, using the fact that the X in these signals has two prongs. The third one defines a more model-independent discovery region, where the X is not assumed to be two-pronged, but rather to be anomalous with respect to the dominant diffuse QCD multi-jet background.

Two-prong regions

In order to provide the greatest possible sensitivity to the $X \to q\bar{q}$ decay, the 2-prong substructure of the X large-R jet candidate is exploited to obtain the strongest possible limits on the cross-section of the model used for interpretation.
The D_2 energy correlation substructure variable (see Chapter 3.4.5) can be used to distinguish jets with two-prong substructure, but when computed with all jet constituents, it is not centrally calibrated. While some W/Z boson calibrated taggers exist, they use the mass information of the boson, rendering them unsuitable for the scenario where the two-pronged boson of interest can have a range of masses. To get around the problem, in this analysis a modified D_2^{trk} is used, that is still an energy correlation functions ratio, but calculated using only tracks information in the jet. In this way, central track uncertainties can be propagated to the D_2^{trk} distribution and therefore used as uncertainties in the analysis.

D_2^{trk} is not meant to be more sensitive to 2-prong jets, D_2 is the "industry standard" and its performances are not expected to be exceeded. A comparison of D_2^{trk} with the standard D_2 variable at analysis preselection can be seen in Figure 4.6. A small systematic shift is observed across the two distributions but they remain close enough to ensure that D_2^{trk} is still sensitive to 2-prong signals with respect to multi-jet background, as demonstrated by signal vs. data plots at preselection in Figure 4.7. As can be seen from this plot, the D_2^{trk} values for the more boosted signal points (m_Y=2000 GeV, m_X=300 GeV) and (m_Y= 3400 GeV, m_X=110 GeV) are lower than the values of background jets from preselection data, as well as of the resolved signal point (m_Y= 5000 GeV, m_X= 2500 GeV).

A custom working point is developed on this variable by examining the optimal cut on D_2^{trk} to maximize the sensitivity as given by:

$$\sigma = \frac{\epsilon}{a/2 + \sqrt{B}} \tag{4.1}$$

where ϵ is the signal efficiency, a represents the number of sigmas corresponding to one-sided Gaussian tests and is set to 3σ, and B is the background yield after the selection. This sensitivity metric comes from [103], where the signal efficiency and the absolute background yields are combined to give the minimum cross section that could be excluded at 95% confidence level.
The background estimation in these studies is data taken from the Higgs mass sideband and normalized to the expected yield taken as the data yield in the unbinned signal region. Since the efficiency of substructure variables is correlated to the boost of the large-R jet and therefore its p_T, these optimal cut plots are remade in bins of p_T to assess the potential for a sliding cut.
Figure 4.8 gives examples of such optimal plot cuts, where the data distribution in a given p_T bin is compared to the combined signal shape. The bottom panel indicates the sensitivity that would be delivered by an upper bound cut at that value of D_2^{trk}. In each of these plots, all boosted signal mass points in the analysis are combined before the optimized selection is determined.

Figure 4.9 shows a summary of the optimal D_2^{trk} cut, combining the results determined in each X p_T bin. Since there is no significant variation in optimal cut value across the p_T spectrum, considering that the last $p_T$bins fluctuations are due to less statistic in those regions, a flat cut of $D_2^{trk} < 1.2$ has been selected.
The Two-prong merged region is therefore defined by the $D_2^{trk} < 1.2$ cut on the X candidate large-R jet.

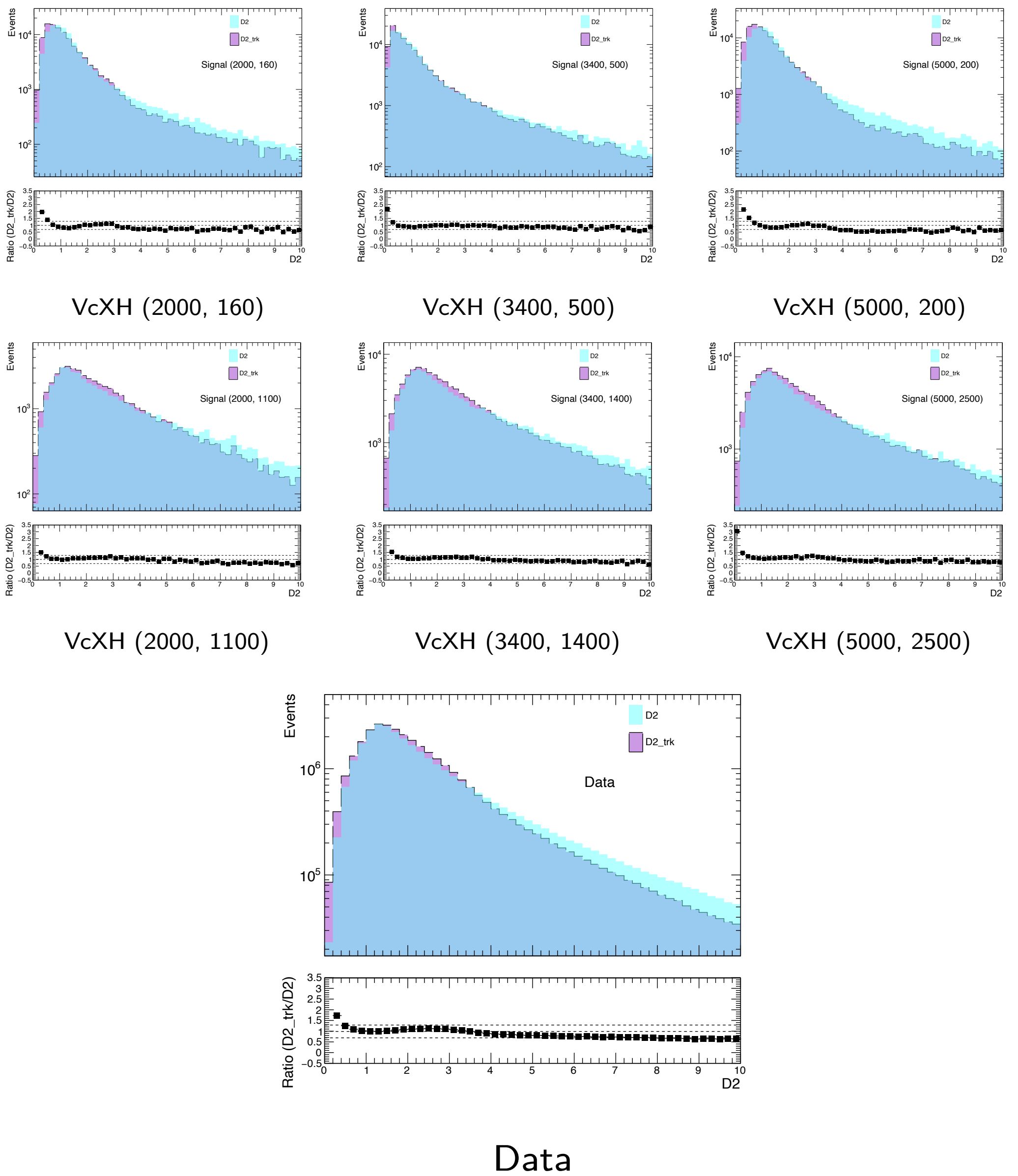


Figure 4.6: Comparison of the standard D2 to the modified D_2^{trk} used in this analysis, using the X candidate large-R jet for several signal points and data at analysis preselection. All distributions are normalized to unity.

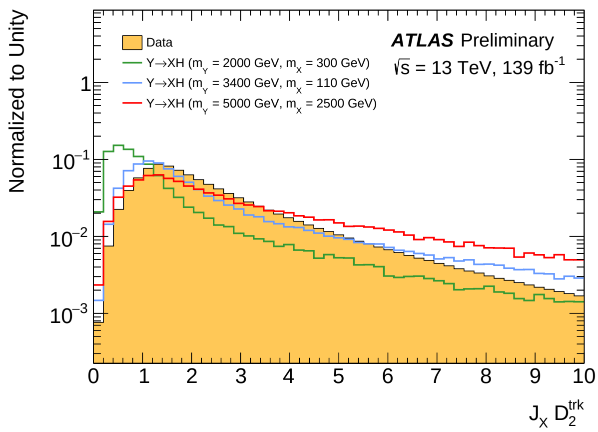


Figure 4.7: Distributions of X candidate D_2^{trk} in data after preselection requirements are applied. Also shown are three $Y \to XH$ simulated signals, labelled by the masses of the Y and X particles. [7]

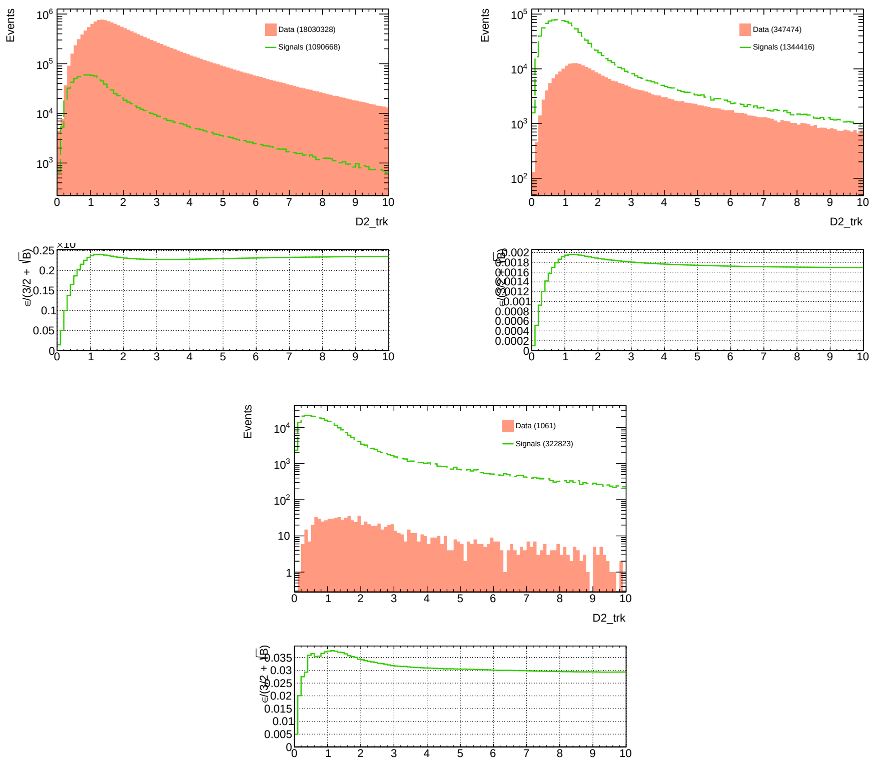


Figure 4.8: Distributions of D_2^{trk} in data and combined $Y \to XH$ signals in bins of p_T (left [500, 700] GeV, center [1100, 1300] GeV, right [2200, 2500] GeV.

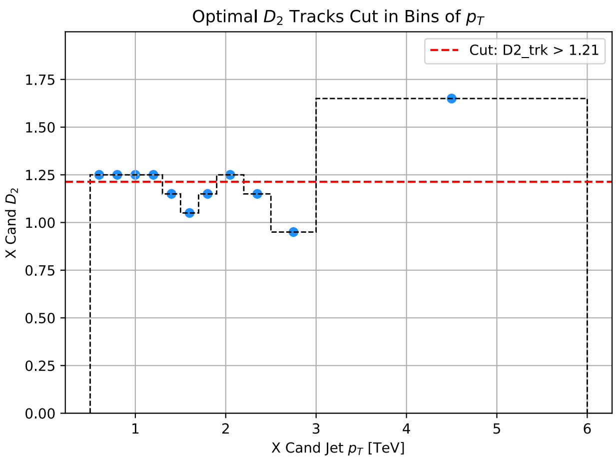


Figure 4.9: A summary of optimal cuts on D_2^{trk}, combining all signal points, in bins of p_T. The last $p_T$bins are affected by greater statistical fluctuations since they are less populated. A flat cut of $D_2^{trk} < 1.2$ (red dashed line; chosen by hand) is used for the X-tagging in the analysis.

For large values of m_X/m_Y, the decay products of the X resonance are no longer much collimated and the large-R jet reconstruction becomes less accurate. The cut on X D_2^{trk}, optimized to maximize the signal sensitivity in merged category, is a good discriminant to reject not-well reconstructed large-R jets since D_2^{trk} has larger values for samples with $m_X/m_Y >\sim 0.3$, as shown in Figure 4.10.
In order to recover sensitivity for these samples not passing the $D2_{trk} < 1.2$ cut

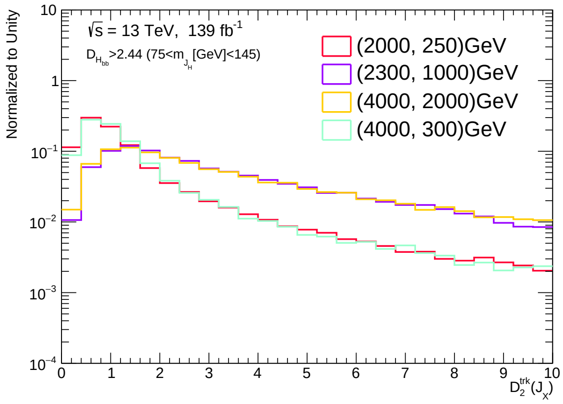


Figure 4.10: D_2^{trk} of the X candidate, for different (m_Y, m_X) signal samples. The distribution spreads over larger values of D_2^{trk} as m_X/m_Y increases.

on the X candidate large-R jet, the X boson is reconstructed via small-R jets, while Higgs boson is still selected as large-R jet.
The assignment of two small-R jets to X is performed with the following algorithm:

1. at least 4 small radius jets are required in the event;
2. small jet pair with the minimum ΔR from the Higgs candidate (reconstructed as large-R jet) is discarded;
3. X boson is reconstructed selecting leading and sub-leading p_T small jets in the remaining jet collection.

Figure 4.11 shows a comparison of reconstructed X and Y masses, for different signal samples with nominal $m_Y = 4000$ GeV, both for events passing and not the $D2_{trk} < 1.2$ cut. m_{XH}^{res} and m_X^{res} are defined as the invariant mass of XH system and X boson respectively, reconstructed using small-R jets for X boson in the region defined by $D2_{trk} > 1.2$.
It is evident that masses in two-prong merged signal region are not well reconstructed when m_X/m_Y increases, while by using the small-R jets for X candidate,

in the orthogonal region, the mass reconstruction capability is preserved.

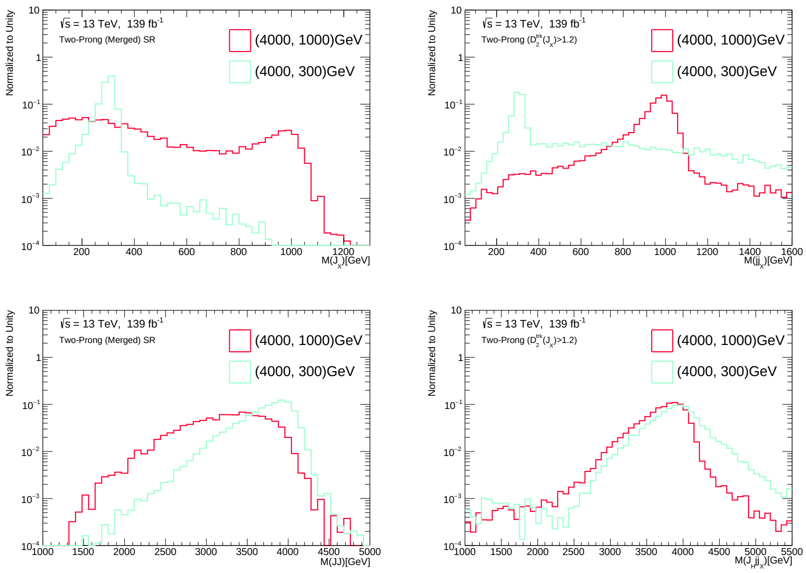


Figure 4.11: Reconstructed X (top) and Y (bottom) masses for (m_Y, m_X) signal samples (4000, 300) GeV and (4000, 1000) GeV, both for $D_2^{trk}< 1.2$ (right) and > 1.2 using small-R jets for X candidate for this latter case (left). It is evident that small-R jets allow to preserve the X candidate mass resolution in the $D_2^{trk}> 1.2$ region.

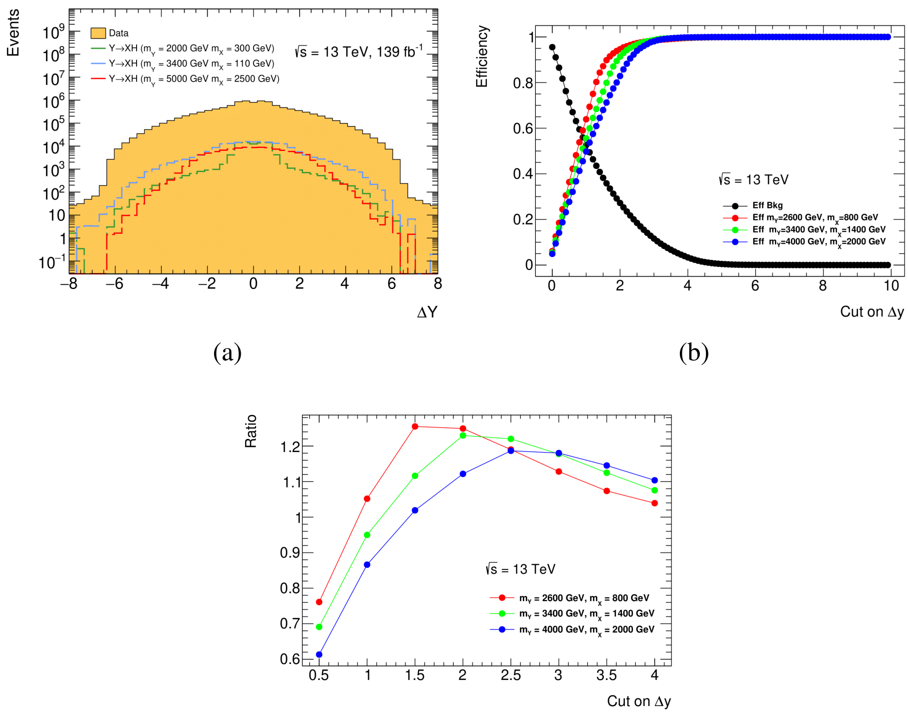
In order to enhance signal sensitivity, the difference in rapidity between the two small jets (ΔY) associated to the X candidate has been studied, since it has a different shape for signals with respect to QCD events, as shown in Figure 4.12. By maximizing the asymptotic significance [70] defined in 4.2, a cut on the absolute value of this variable to be $<$ 2.5 has been chosen.

$$Z = \sqrt{\sum_{i}^{N} 2 \cdot [(S_i + B_i) \cdot log\,(1 + S_i/B_i) - S_i]} \tag{4.2}$$

An additional selection on the p_T balance, defined as

$$(j1\ p_T - j2\ p_T)/(j1\ p_T + j2\ p_T)$$

of the two small jets ($j1$ and $j2$) associated to X has been studied, since this variables shows remaining discrimination power after the cut on ΔY is applied, as can be seen in Figure 4.13.

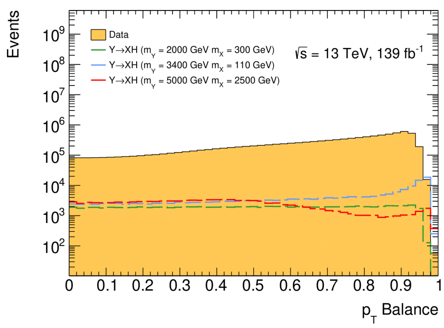


(a) (b)

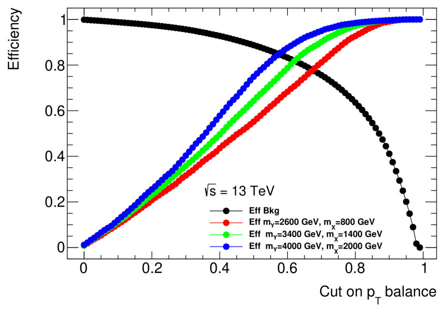
(c)

Figure 4.12: Rapidity difference $\Delta\ Y$ for the two selected small jets (a); signals efficiency (ϵ_S) and background rejection (1 - ϵ_{QCD}), varying the cut on $|\Delta\ Y|$ (b); ratio between the asymptotic significance calculated after and before applying a varying cut on $|\Delta\ Y|$ (c). Tree signals are shown: (m_Y=2600 GeV, m_X=800 GeV), (m_Y=3400 GeV, m_X=1400 GeV) and (m_Y=4000 GeV, m_X=2000 GeV).

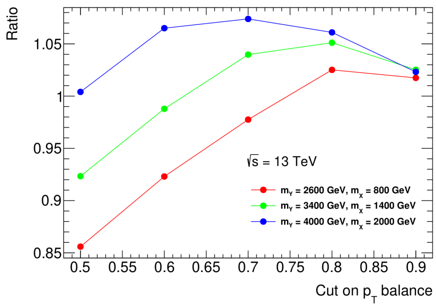
The chosen cut on pT balance is < 0.8.
The final two-prong resolved selection is defined by the cuts: $D_2^{trk} > 1.2$ AND $|\Delta Y| < 2.5$ AND p_T balance < 0.8.



(a)



(b)



(c)

Figure 4.13: Small jets p_T balance (a); signals efficiency (ϵ_S) and background rejection (1 - ϵ_{QCD}), varying the cut on p_T balance (b); ratio between the asymptotic significance calculated after and before applying a varying cut on p_T balance (c). The cut $|\Delta Y| < 2.5$ is already applied. Tree signals are shown: (m_Y=2600 GeV, m_X=800 GeV), (m_Y=3400 GeV, m_X=1400 GeV) and (m_Y=4000 GeV, m_X=2000 GeV).

An overview of the gain introduced by the resolved selection in the (m_X,m_Y) plan is shown in Figure 4.14, where for each signal point the ratio between the cross section at 3σ (asymptotic) significance obtained in merged region and that obtained by summing in square significance in merged and resolved are calculated. Higher values correspond to lower combined limits. The significance gain starts to become evident for m_X/$m_Y$$\geq$ 0.3.

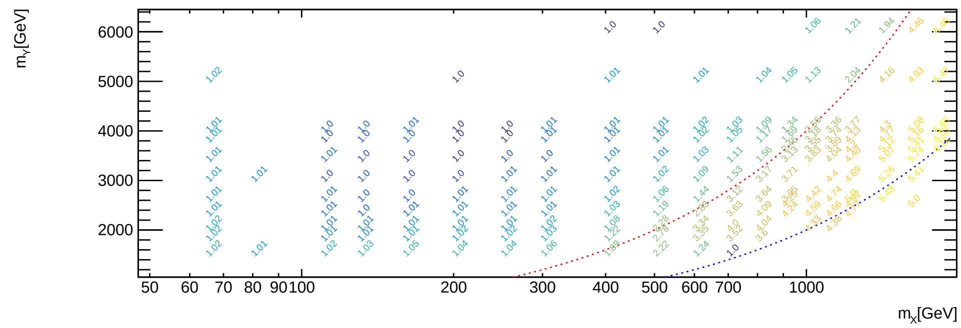


Figure 4.14: Ratio between the cross section at 3σ asymptotic significance calculated in two-prong merged region and that calculated combining merged and resolved regions (by summing in quadrature the respective significances) for each simulated signal point. The red dotted line corresponds to m_X/m_Y=0.3, the blue dotted line to m_X/m_Y=0.5.

Anomaly region

To mitigate model-dependence, an additional region is built that relies on a fully data-driven *anomaly score (AS)*. The AS is defined using a variational recurrent neural network (VRNN) [48] , which consists of a variational autoencoder (VAE) whose latent space is updated at each time step of a recurrent neural network (RNN). This architecture combines the variational inference capabilities of a VAE with the sequence modeling provided by an RNN.
A first application of this methodology in a dijet search context was performed using the LHC Olympics simulated multijet dataset [88].

A standard autoencoder has two stages. The first encoder step reduces the dimensionality of an input by determining its most salient descriptive features and encodes it into a latent space. The second decoder step samples from that latent space and attempts to construct an output of the original dimensionality and as similar as possible to the input. The loss in the training is often therefore based on the reconstruction error of output to input. The resulting net has thus learned the most important features for describing elements in a training set.
An autoencoder can be enhanced into a variational autoencoder (VAE), which performs Bayesian inference by sampling from a multivariate Gaussian latent space in the decoder stage.
Finally, a VRNN is an RNN with an autoencoder at each training time step, where the updating of the hidden state during the RNN training corresponds to updating the VAE latent space. Figure 4.15 shows a diagram of a VRNN cell.

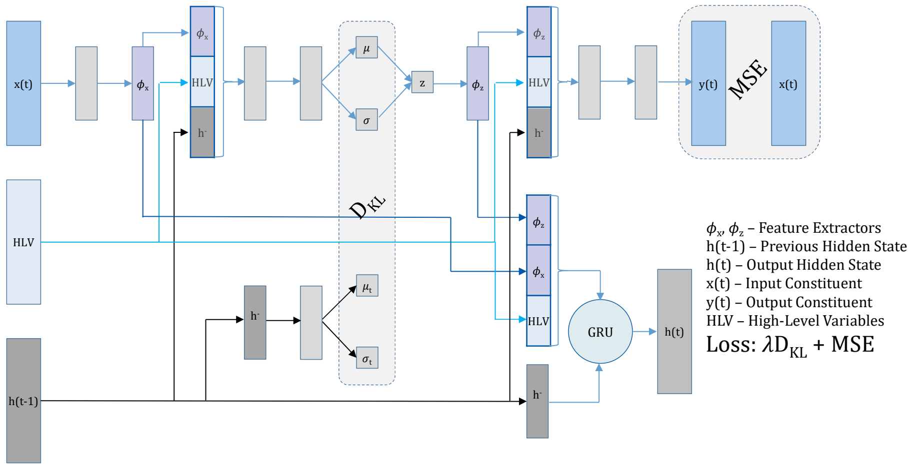


Figure 4.15: Diagram of a VRNN cell, with the autoencoder portion at the top of the image and the RNN architecture at the bottom.

The VRNN is applied in this analysis through its output anomaly score on the X candidate large-R jet. AS is used as discriminating variable to construct a region enriched in anomalous jets (i.e. X candidates) and depleted in multijet background.

The VRNN is trained over large-R TCC jets in the ATLAS Run 2 dataset satisfying the trigger plateau criteria, the jet requirements described in Section 3, and $p_T > 1.2$ TeV. The p_T selection is designed to restrict input jets to highly boosted topologies, which are both well-described by the k_t-sorted sequence modeling and the most difficult to distinguish with regular substructure methods.
90% of the data is used for training and 10% for validation. As the input consists solely of jets from data, no labeling scheme is used in training, distinguishing this method of unsupervised learning from traditional supervised machine learning where the input is labeled according to a signal or background categorization.

The input jets are modeled as a sequence of up to twenty constituent four-vectors per jet, ordered in k_t splitting starting from the highest p_T constituent. The training is also conditioned over four high-level variables, namely the energy correlator substructure variable D_2 [95] for two-prong sensitivity, the N-subjettiness ratio

τ_{32} [109] for three-prong sensitivity, and the two k_t-splitting scale ratios d_{12} and d_{23} [98]. This input modeling is designed to reveal correlations between constituents and substructure, allowing the VRNN to distinguish jets with anomalous energy deposition patterns from the background of homogeneous jets originating from QCD processes.
An alignment procedure is applied to each jet that re-scales to the same p_T, boosts to the same energy, and rotates to the same orientation in η and ϕ, minimizing the ability of the VRNN to tag only on anomalous kinematic properties without considering internal constituent properties. Additionally, the mass re-scaling is performed by providing an overall multiplicative factor for each constituent and the boosting is performed by Lorentz boosting each constituent along the jet's axis by the same boost factor.
The high-level variables are fed to the network as a separate input vector from the constituent sequence, and is accessed by the model in each time-step. In more detail, the high-level variables input vector is concatenated with contextual feature-extracting layers and the hidden state before being accessed by the decoder, encoder, and hidden-state updating architectures in the VRNN cell. This process can be seen graphically in Figure 4.15.

The AS is derived from the VRNN loss (Equation 4.3) which is composed of two terms: a reconstruction error term, to minimize differences of the decoded result with respect to the original input, and a Kullback-Leibler (KL) divergence term, to constrain the spread of probability classes in the latent space.

$$\mathcal{L}(t) = |\mathbf{y}(t) - \mathbf{x}(t)|^2 + \lambda D_{KL}(z||z_t). \tag{4.3}$$

Here, x represents the input vector, y the output, t is the current time step, z is the approximate posterior, z_t is the learned prior, and λ is a hyperparameter that weights the KL divergence term in importance. A loss with both of these terms ensures correct functionality of the VRNN.
Minimizing the reconstruction loss ensures that the VRNN correctly characterizes and learns relevant features of the input dataset. Adding the KL divergence term punishes the VRNN for using too many dimensions in the input latent space to describe the input, ensuring that only the most important features are used and overly precise classification does not occur.
An overall loss $\mathcal{L}$ over the sequence is then computed by averaging the individual time-step losses over the length of the sequence N:

$$\mathcal{L} = \frac{\mathcal{L}(t)}{N}. \tag{4.4}$$

The resulting per-jet anomaly score ρ is a simple function of the KL divergence term, as shown in Equation 4.5.

$$\rho = 1 - e^{-\overline{D_{KL}}}. \tag{4.5}$$

Scores used in the analysis to select the X candidate are further subject to a transformation (Equation 4.6) that ensures the mean of the score distribution is 0.5 and more anomalous jets populate higher values of the anomaly score.

$$\rho' = 1 - \left(\frac{\rho}{2\overline{\rho}}\right), \tag{4.6}$$

A custom architecture was developed and trained using the PyTorch deep learning library [?]. A scan has been performed of various hyperparameter options to ensure the optimal values are used in the training, giving the highest area under curve (AUC) after testing across all signal models. The VRNN hyperparameters are the number of dimensions in the hidden state (hdim), the number of dimensions in the sampled latent variables (zdim), and the kl-weight. The chosen combination is: hdim = 16, zdim = 2 and kl-weight = 0.1
The network is updated using the Adam optimizer with a learning rate parameter of 10^{-5}. No regularization via weight decay is applied, however gradient clipping is implemented with a clip value of 10. Training is performed in batches of 256 jets per batch, and continues until 100 training epochs have been completed, at which point the model reaches a performance plateau.

Figure 4.16 shows the resulting anomaly score of the X candidate, in preselected data and representative signals. No $D_{H_{bb}}$-tagging is applied, therefore these regions are more background-rich than anticipated in the final signal region, and enhancement of the signal over data is expected to increase substantially. Both $Y \to XH$ signals are shown, with a 2-prong X hypothesis, and so-called ”special” signals, with varying X candidate hypotheses. Specifically we examine heavy-flavor, 3-prong, and dark jet samples, to vet the capability of the anomaly in providing broad model sensitivity.

The optimal cut on the anomaly score for the analysis is found using a similar procedure to the one used for D_2^{trk}. The sensitivity σ is scanned as a function of various cuts on the X candidate AS for signals combined together and in bins of large-R jet p_T.
Figure 4.17 shows a summary of the optimal AS cut, combining all $Y \to XH$ signal points and binning in X candidate jet p_T. As with D_2^{trk}, no significant correlation with p_T is present and a flat cut of AS > 0.5 is placed on the X candidate.

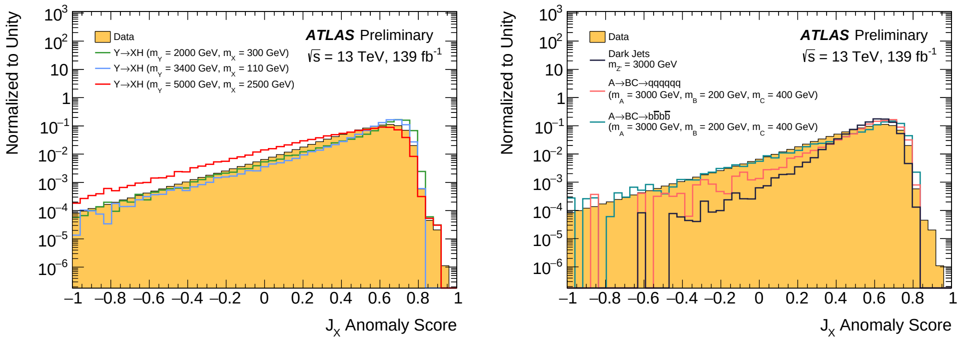


Figure 4.16: Distributions of the X candidate anomaly score (AS) in data after preselection requirements are applied. Also shown are three $Y \to XH$ simulated signals (left), labelled by the masses of the Y and X particles, and the three additional signals with alternative X decay hypotheses, namely heavy flavor, three-prong, and dark jet (right). All distributions are normalized to unity. [7]

The AS as a variable has a range of -1 to 1; the fact that the optimal cut is at –1 for higher pT means that the best option for S/B discrimination is no AS selection at all. This is attributed to the fact that background drops off rapidly at high pT (> 2 TeV). The flat cut at 0.5 is designed to be optimal for the lower pT range, and maintain a very low background search region for higher pT (so that signal is still prominent.)
The third signal region, the Anomaly, is therefore defined by applying the cut $AS > 0.5$ on the X candidate large-R jet, instead of the cut on D_2^{trk}.

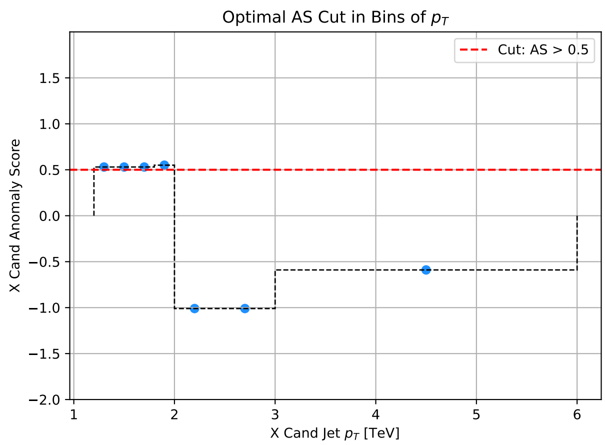


Figure 4.17: A summary of optimal cuts on anomaly score, combining all signal points, in bins of p_T. Negative values are due to few statistics.

4.3.4 H candidate selection

A selection on the Higgs boson candidate variables is performed after the X selection is applied, to sort the analysis into three categories. For all signal regions, a working point that provides a flat 60% efficiency across jet p_T is applied ($D_{H_{bb}} > 2.44$) to the Higgs boson candidate, along with a mass window requirement of 75 GeV $< m_H <$ 145 GeV. As a result, the background in the signal region is expected to be overwhelmingly QCD multi-jet processes, with a few percent contamination from top or V+jets processes, therefore Higgs Mass sidebands are used for the data-driven background derivation (see Appendix B.1).

In total, three signal regions (*two-prong merged*, *two-prong resolved* and *anomaly*) and 15 background estimation regions (5 non-SR regions for each SR) are defined. A summary of the preselection cuts and the final regions definition can be found in Table 4.2.

Parameter	Preselection requirements
m_{JJ} [GeV]	> 1300
$p_{\mathrm{T}}(J_1)$ [GeV]	> 500
m_J [GeV]	$m_{J_1} > 50$ OR $m_{J_2} > 50$
$D_{H_{bb}}$	$>$ -2

	Signal regions		
	Two-prong merged	Two-prong resolved	Anomaly
[GeV]	(75, 145)		
$D_{H_{bb}}$	> 2.44		
D_2^{trk}	< 1.2	> 1.2	-
$\lvert\Delta y_{j_1,j_2}\rvert$	-	< 2.5	-
p_{T}^{bal}	-	< 0.8	-
Anomaly Score	-	-	> 0.5

	Background estimation regions				
	CR0	HSB0	HSB1	LSB0	LSB1
[GeV]	(75, 145)	(145, 200)		(65, 75)	
$D_{H_{bb}}$	< 2.44	< 2.44	> 2.44	< 2.44	> 2.44

Table 4.2: Specific selections defining the SRs and background estimation regions. $J_{1(2)}$ and $j_{1(2)}$ are respectively the p_T-(sub)leading large-R and small-R jets.

Each of these regions is actually further split according to m_X value, since the signal search in the two-dimensional space of m_Y versus m_X employs sliding windows of the X candidate mass spectrum, dividing the data into a series of overlapping m_X ranges for which the m_{JJ} distribution is fit (see Section 4.7).
Figure 4.18 shows an illustration of the selection flow, along with the analysis regions based on H candidate cuts.
Acceptance times efficiency of the entire $Y \to XH$ signal grid in the merged and resolved SRs is shwon in Figure 4.19.

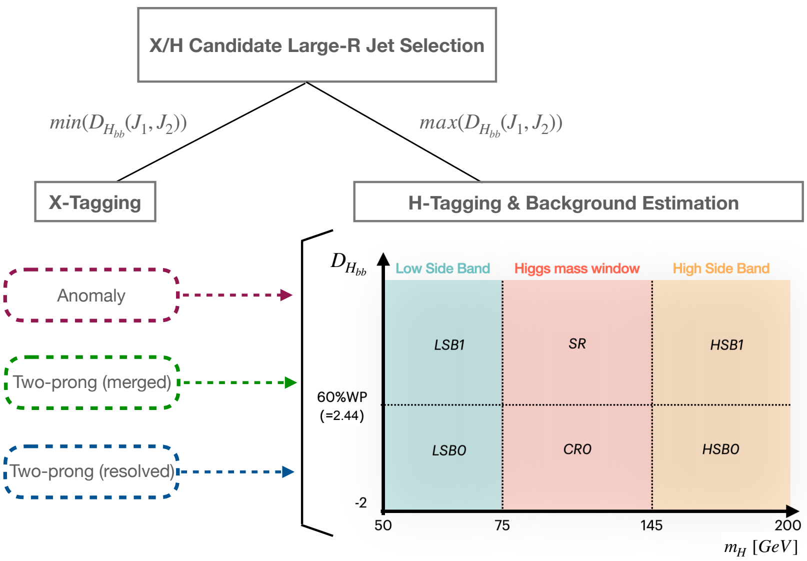


Figure 4.18: Illustration of the selection flow after preselection and the analysis regions of the $Y \to XH$ search. Preselection events are sort into three separate categories, namely two-prong merged, two-prong resolved and anomaly. Each region (LSB0, LSB1, CR0, SR, HSB0 and HSB1) is defined in the same way for all three analyses using the H candidate mass and $D_{H_{bb}}$ score. [7]

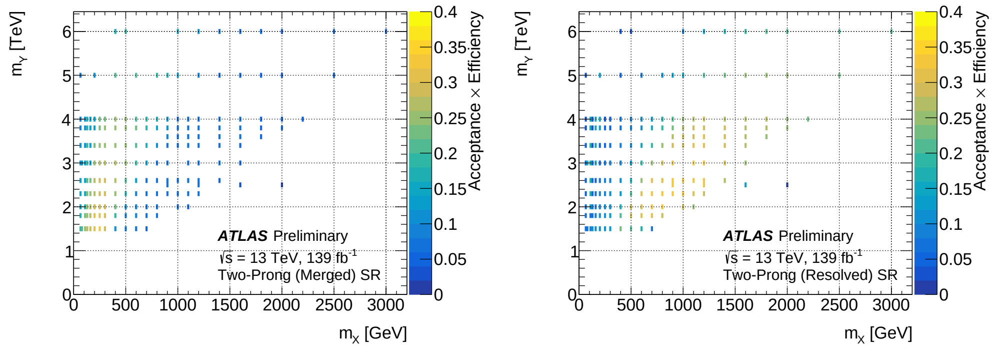


Figure 4.19: Acceptance $\times$ efficiency for every signal point at 139 fb^{-1} luminosity in the two-prong merged and resolved SRs. Each signal point is assumed to have a cross section of 1 pb. [7]

4.4 Background Estimation

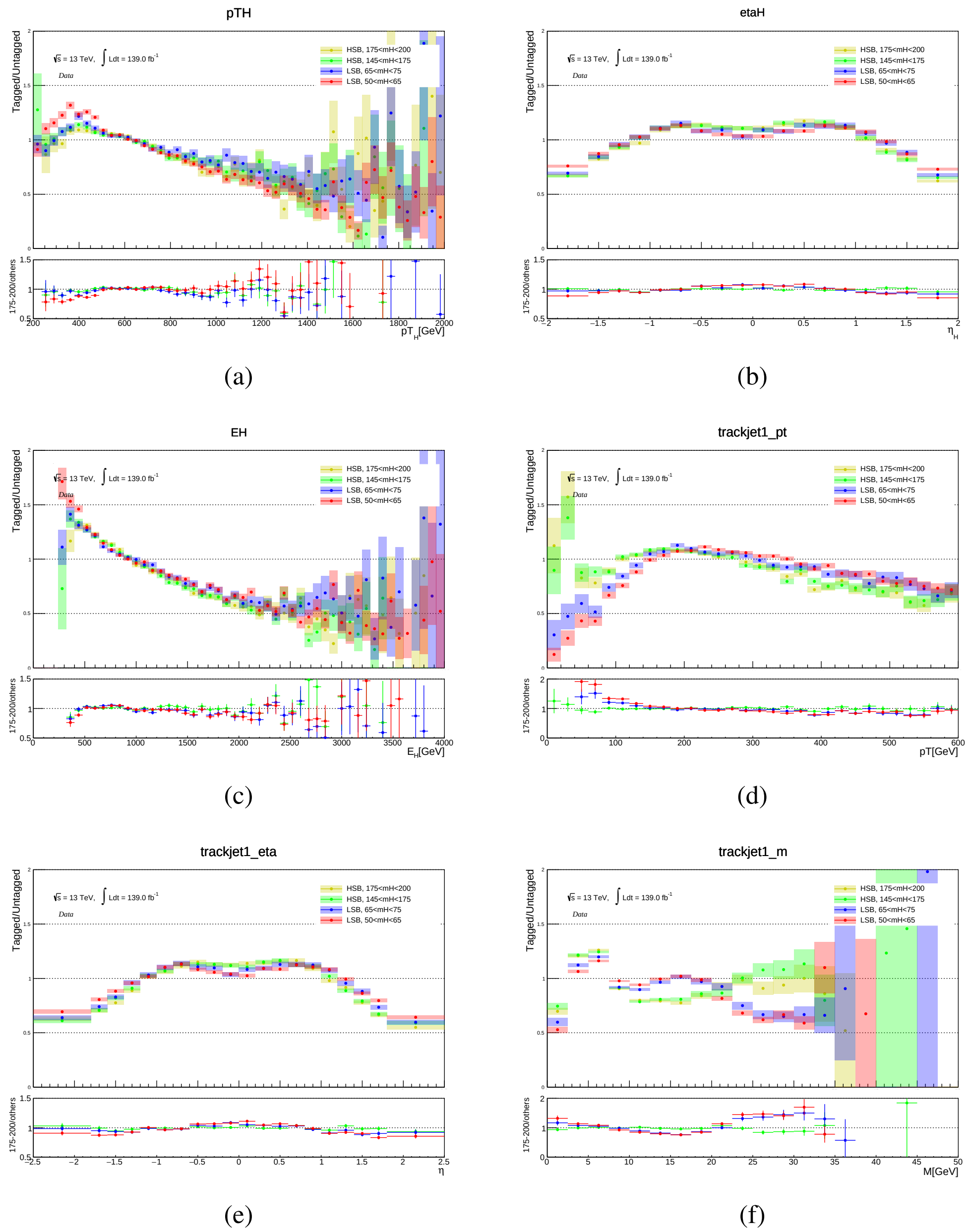
The background in the SRs mainly arises from high p_T multi-jet events (more details on the background composition are shown in Appendix B.1). Simulation for such QCD processes includes well-known mismodelings and is computationally expensive to generate. Therefore in this analysis, the background estimation is fully data-driven and derived from regions that are orthogonal to the SR based on the Higgs boson jet criteria.

The shape of the expected m_{JJ} distribution in the SR is obtained from data in the CR0 and a reweighting procedure is necessary to account for the effect of $H \to b\bar{b}$ tagging. This procedure is developed in the HSB and validated by applying the weights to data in LSB0 and comparing the resulting m_{JJ} spectrum to that observed in LSB1 data. The generation of weights is performed inclusively in the X candidate selection and applied separately to create three background predictions, one for each SR.

The baseline for this method is the assumption of independence of the reweighting function from the Higgs candidate mass, in such a way that it is possible to define it in a certain mass window and then apply it in another window, not far from the first one.
Figures 4.20 and 4.21 show normalized data shapes, compared between $D_{H_{bb}}$-tagged and -untagged regions, for those variables more sensitive to the $D_{H_{bb}}$efficiency cut, among the following Higgs mass windows:

- $mH \in (50, 65)$ GeV
- $mH \in (65, 75)$ GeV
- $mH \in (145, 175)$ GeV
- $mH \in (175, 200)$ GeV

For most of these variables the independence of $D_{H_{bb}}$from m_H can be considered valid within experimental errors. The most discrepant region is $m_H \in (50, 65)$ GeV, closer to the lower cut on large-R jet mass, therefore this region is not used to validate the method.



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

Figure 4.20: Ratios between $D_{H_{bb}}$-tagged and -untagged regions, across m_H windows, for (a) p_T, (b) η and (c) E of the Higgs candidate jet; (d) p_T, (e) η and (f) m of the p_T-leading track jet associated to Higgs candidate large-R jet.

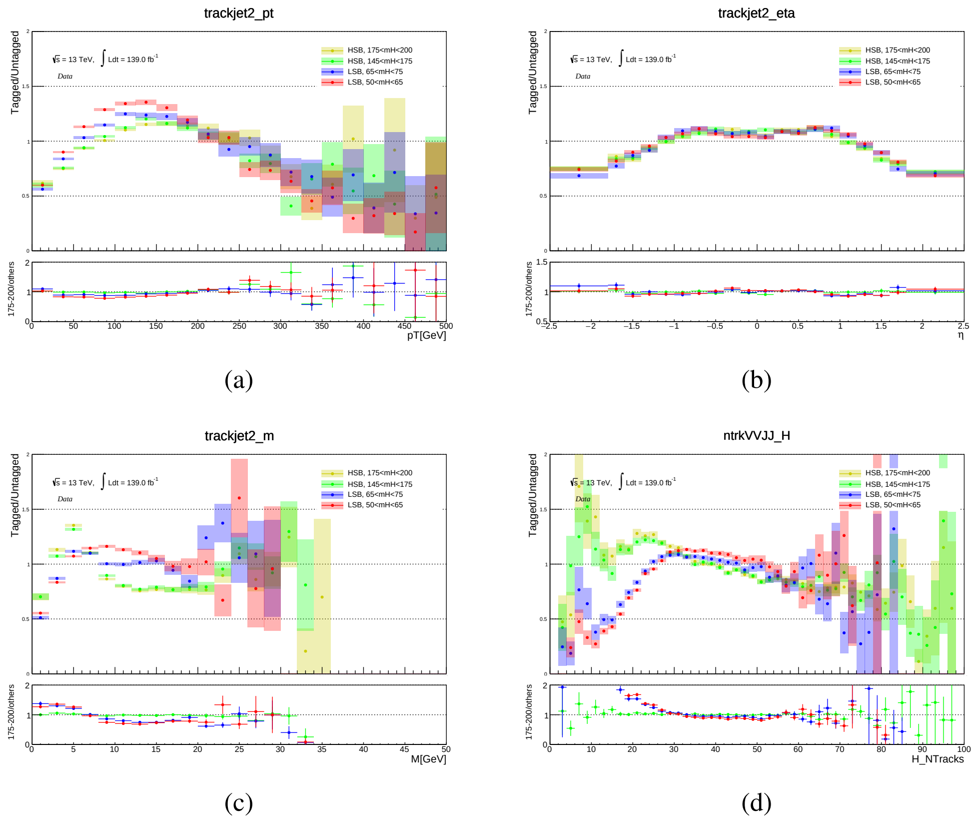


Figure 4.21: Ratios between $D_{H_{bb}}$-tagged and -untagged regions, across m_H windows, for (a) p_T, (b) η and (c) m of the p_T-subleading track jet associated to Higgs candidate large-R jet and (d) the number of tracks associated to the Higgs candidate large-R jet.

4.4.1 The reweighting problem with a Machine Learning algorithm

The problem of defining a reweigthing function to map the kinematic of a certain region into another can be transformed into the problem of estimating a likelihood ratio between the two regions.
If p_0 and p_1 are the probability distribution functions (*pdf*s) of a multi-dimensional variable X, sampled from two kinematic regions labelled 0 and 1 respectively, the needed reweighting factor corresponds indeed to their *Likelihood Ratio* (or *Importance*) $w(X)$:

$$p_1(X) = w(X) \cdot p_0(X) \rightarrow w(X) = \frac{p_1(X)}{p_0(X)}. \tag{4.7}$$

In this analysis context, p_0 and p_1 would respectively be the pdfs of CR0 SR, and X the set of all the possible kinematic variables associated to each event.
The computation of the likelihood ratio relies on the knowledge of the probability density functions (*PDF*s) which, for the majority of the cases, it's not simple and straightforward to obtain. However, if the final purpose is only building the likelihood ratio, it is possible to bypass this step and to perform the *Direct Importance Estimation*, where the PDFs ratio is the quantity directly calculated from data, without the need of knowing the single likelihoods functions.

The procedure adopted in the previous round of the analysis is to roughly estimate the PDFs by one- or two-dimensional histograms. Reasonable variables to choose to create these histograms are the ones with the greatest differences between the two regions, then the weights map is obtained by dividing them between HSB1 and HSB0. Although this approach yields good results, a dedicated fine-tuning to keep into account correlations among variables is needed and performances get worse on the reweigthing of those variables not much correlated with the ones used in the reweighting procedure.
In this new round of the analysis a new approach based on a Machine Learning technique has been introduced. This method allows a direct estimation of PDF ratios and permits to deal with the intrinsic multi-dimensionality of the Likelihood ratio.

The basic idea is to formulate the problem as an optimization problem to be solved by a Neural Network (NN).
Let's first settle the direct importance estimation problem as a least-squares function fitting problem [89], that can be easily generalized to a more generic one.
From this point on, the estimate of quantities on data will be indicated with the symbol $\widehat{\ }$.

Let's consider two probability density functions p_0 and p_1 and suppose we have N_0 independent and identical distributed (i.i.d.) measurements extracted from p_0 and N_1 i.i.d. measurements extracted from p_1 of B observables: $\{X^0_{i,\,j}\}^{N_0,\,B}_{i=1,\,j=1}$ and $\{X^1_{l,\,j}\}^{N_1,\,B}_{l=1,\,j=1}$.
We know neither the analytical form of $p_0(\underline{X})$ and $p_1(\underline{X})$ nor their empirical realization but we are interested in the ratio between them and we can estimate it from the two samples of the extracted measurements.
Let's call $w(\underline{X}) = \frac{p_1(\underline{X})}{p_0(\underline{X})}$ the analytical ratio and $\hat{w}(\underline{X}) = \sum_f \alpha_f \cdot \phi_f(\underline{X})$ its estimate from data, which can be written as a linear combination of arbitrary chosen functions of data $\phi_f(\underline{X})$ that must satisfy the same properties of $w(\underline{X})$:

- $w(\underline{X}) = \frac{p_1(\underline{X})}{p_0(\underline{X})} > 0$
- $w(\underline{X}) \to +\infty$ if $\underline{X} \in p_1(\underline{X})$
- $w(\underline{X}) \to 0$ if $\underline{X} \in p_0(\underline{X})$

The coefficients $\underline{\alpha}$ are constants that are chosen in a way to obtain the best estimate of $w(\underline{X})$, by minimizing a loss function which depends on them.

Now that we have fixed the notation we can formulate our minimization problem by defining a loss function as the sum of the quadratic differences between the true value $w(\underline{X})$ and its estimates on data $\hat{w}(\underline{X})$.
Thus the loss function J_0 can be written with the following expression:

$$\begin{aligned} J_0(\underline{\alpha}) = E_0[(\hat{w}(\underline{X}) - w(\underline{X}))^2] = \int d\underline{X}\; p_0(\underline{X})\; (\hat{w}(\underline{X}) - w(\underline{X}))^2 = \\ \int d\underline{X}\; p_0(\underline{X})\; \hat{w}(\underline{X})^2 \;+ \int d\underline{X}\; p_0(\underline{X})\; w(\underline{X})^2 \;-\; 2\int d\underline{X}\; p_0(\underline{X})\; \hat{w}(\underline{X}) \cdot w(\underline{X}) = \\ = \int d\underline{X}\; p_0(\underline{X})\; \hat{w}(\underline{X})^2 \;- 2 \cdot \int d\underline{X}\; p_1(\underline{X})\; \hat{w}(\underline{X}) \end{aligned} \tag{4.8}$$

where J_0 depends on $\underline{\alpha}$ because $\hat{w}(\underline{X})$ depend on it.
In the expression in the centre of 4.8 the second term does not depend on data thus it can be ignored and the expression can be written as follow:

$$J_0(\underline{\alpha}) = E_0[\hat{w}(\underline{X})^2 - 2 \cdot w(\underline{X}) \cdot \hat{w}(\underline{X})] = E_0[\hat{w}(\underline{X})^2] - 2 \cdot E_1[\hat{w}(\underline{X})] \tag{4.9}$$

This loss function has to be minimized with respect to the constants $\underline{\alpha}$ on data, in order to obtain the best estimate of $w(\underline{X})$.
If $\hat{w}(\underline{X})$ is intended as the output of a Neural Network (NN) and the constants $\underline{\alpha}$

as its weights $\underline{\theta}$, the loss function J_0 can effectively be intended as a loss function to be minimized by a NN.

Since $\hat{w}(\underline{X})$ is positively defined, a constraint term should be added to the loss function, or the output can be transformed with a monotone function and the loss consequently adapted. For example, choice of log-likelihood ratio estimation offers a straightforward solution to the problem of non-negative likelihood ratio. Applying this transformation the network output will be an estimate of $\ln(w(\underline{X}))$. A more generic way to write the loss function for a function of $u(\underline{X}, \underline{\theta})$ can be the following [101]:

$$J_0(\underline{\theta}) = E_0[\phi(u(\underline{X}, \underline{\theta})) + w(\underline{X}) \cdot \psi(u(\underline{X}, \underline{\theta}))] = E_0[\phi(u(\underline{X}, \underline{\theta}))] + E_1[\psi(u(\underline{X}, \underline{\theta}))] \tag{4.10}$$

Here ϕ and ψ are two real functions of real variables that should satisfy some mathematical criteria if we want $u(\underline{X}, \underline{\theta})$ to minimize the loss function.
It can be demonstrated that there are not single solutions giving a closed form for ϕ and ψ. A possible choice for $u = \ln(w)$ can be $\phi(u) = \sqrt{e^u}$ and $\psi(u) = \frac{1}{\sqrt{e^u}}$.
With these substitutions, the loss function to minimize becomes:

$$J_0(\underline{\theta}) = E_0[\sqrt{e^{u(\underline{X}, \underline{\theta})}}] \ + \ E_1[\frac{1}{\sqrt{e^{u(\underline{X}, \underline{\theta})}}}] \tag{4.11}$$

A Deep Neural Network has been used to minimize the loss in equation 4.11, in the training phase, in order to estimate the weights for mapping HSB0 into HSB1 and to obtain a neural network model able to predict weights between CR0 and SR.
The training has been performed in the HSB, using data both in tagged (HSB1) and untagged regions (HSB0), where the Higgs candidate mass is required to be greater than 145 GeV and less then 175 GeV. Then the DNN is read on the CR0 events, in order to obtain the background shape in SR. The chosen training region does not cover the whole HSB because the region with higher Higgs mass (called *alternative training region*) is used for systematic uncertainties estimation, as it will be subsequently explained in section 4.6.3. The estimated background is validated in the LSB, only in the window defined by 65 GeV $< m_H <$ 75 GeV, because the assumption of $D_{H_{bb}}$being independent from m_H is no longer valid below 65 GeV, as explained in 4.4.3.

Table 4.3 contains the amount of data on each of the dataset used in the training and validation procedure. It can be seen that in both training and alternative training region there are almost the same data, as it must be to not affect the estimate of the systematics with differences in the dataset size. The proportion tagged/untagged is around 0.06 and it is stable across m_H. Also data for untagged

	full HSB $mH \in [145, 200]GeV$	train region $m_H \in [145, 175]GeV$	alternative train region $m_H \in [165, 200]GeV$	test region $m_H \in [65, 75]GeV$
tag: $D_{H_{bb}}$>2.46	87107	52647	50422	39167
ϵ wrt full tag HSB	1.00	0.60	0.58	0.45
untag: (-2)< $D_{H_{bb}}$< 2.46	1396700	844284	812570	621082
ϵ wrt full untag HSB	1.00	0.60	0.58	0.44
untag without lower $D_{H_{bb}}$bound: $D_{H_{bb}}$<2.46	3079960	1862210	1789210	1432470
ϵ wrt full untag HSB (no $D_{H_{bb}}$cut)	1.00	0.60	0.58	0.47
tag/untag	0.062	0.062	0.062	0.063
tag/untag (no $D_{H_{bb}}$cut)	0.028	0.028	0.028	0.027

Table 4.3: Data in each of the region used for the nominal training, alternative training and testing of the DNN for the background estimation. Also the efficiency (ϵ) with respect to the full HSB regions, in order to look at the relative proportion of each region, and the ratio tagged/untagged are reported. Data for untagged region without cut $D_{H_{bb}}$>(-2) are twice the untagged region adopted.

region without cut $D_{H_{bb}} > -2$ are shown in the table, and it can be seen that with this cut about half of untagged data are removed. To have more statistics in untagged region is not useful for the network training, since it is already a very large sample, while it would insert a bigger imbalance between tagged and untagged categories.

The train is performed after the H/X ambiguity resolution is solved, without any special tagging on X candidate. This inclusive training allows to have a single weights set that is then split for Two-prong and Anomaly regions.
The dataset has been divided in training and test sets using the 70% and 30% of the full training dataset, respectively. From the training set, the 20% is used for validation, in order to validate the model and to monitor the overfitting[2] during the training phase.

The network structure used is a fully connected API Sequential model from Keras [46] with 3 fully connected inner layers, each with 20 neurons and a rectified linear unit activation function (*ReLU*). In order to reduce the problem of the overfitting during training, 10 % of connections among inner layers is randomly truncated (this is called *dropout* and referenced in [107]).
The last layer has a single output with a simple linear activation function, since

[2]The world *overfitting* in machine Learning refers to a situation where a model learns the details and noise in the training data to the extent that it negatively impacts the performance of the model on new data. This means that the noise or random fluctuations in the training data is picked up and learned as concepts by the model. The problem is that these concepts do not apply to new data and negatively impact the models ability to generalize. On the other hand, the *underfitting* refers to a model that can neither model the training data nor generalize to new data. An underfit machine learning model is not a suitable model and will be obvious as it will have poor performance on the training data.

we are interested in the value minimizing the loss, without any other function applied on it. In this way, the network output is exactly an estimate of the quantity $\ln\left(\frac{p_1(\underline{X})}{p_0(\underline{X})}\right)$.
The model is trained using the Adam optimizer in Keras with Tensorflow as backend [15].
The customized loss function in 4.11 has been used. Since in the formula there are the expectation values with respect to p_0 and p_1, they are calculated as the averages on data, where N_0 and N_1 are respectively the number of events belonging to the two categories and y is the event label, equal to 0 for events belonging to p_0 and 1 for events belonging to p_1:

$$Loss = \frac{1}{N_0}\sum(1-y)\sqrt{e^u} + \frac{1}{N_1}\sum y\frac{1}{\sqrt{e^u}} \tag{4.12}$$

Variables given in input to the network are the following:

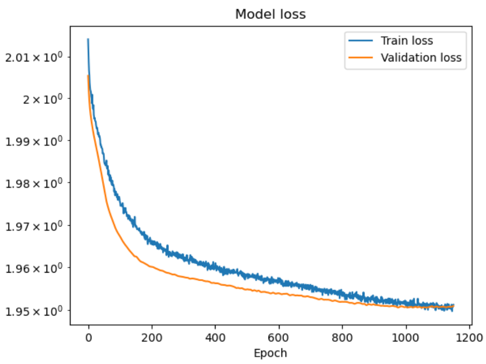
- p_T, η, ϕ, E of the Higgs candidate
- the number of tracks associated to the Higgs candidate
- p_T, η, ϕ, m of the first two track jets associates to the Higgs candidate, ordered in p_T

Each variable is standardized before being fed to the network, with the transformation $x = \dfrac{(x-\mu)}{\sigma}$, where μ and σ are the mean and the standard deviation of the distribution of the x variable.

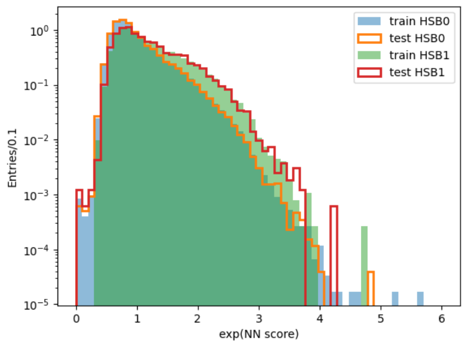
The network is trained only on events with at least two track jets associated to the Higgs candidate, while it is read on all the events. For events with less than two track jets, the average value of missing variables in HSB is used as default. This procedure is safe from inserting bias in the estimated background since the fraction of events with less than 2 track jets associated to Higgs candidate fat jet is really small: it is $< 6\%$ for signals and $< 0.02\%$ for data after pre-selection cuts and the fraction for signal decreases in SR.

Training is performed using a batch size equal to the full dataset size. A comparison with a batch size equal to 2048 has been done; results are identical within the statistical errors. The chosen training epochs are 1600, with early stopping at 100 epochs if the value of the loss calculated on the validation dataset does not increase for 100 subsequent epochs.

In Figure 4.22a training and validation loss functions are shown; Figure 4.22b is the exp of the network output on HSB0 and HSB1, both for training and test dataset.



(a)



(b)

Figure 4.22: (a) Loss functions calculated on training (red curve) and on validation (orange curve) dataset, during network training. (b) The exponent of network predictions, that is an estimate of $\frac{pdf_{\ HSB1}}{pdf_{\ HSB0}}$, both for training and test dataset.

Since the output of the DNN is an estimate on data of the logarithm of the likelihood ratio, the score is transformed by applying the exponential function to it in order to obtain the event-level weight for each event. This method allows to obtain normalized distribution, while the background normalization factor in SR is obtained by the fit procedure.
With the DNN method the full background shape is estimated. In the previous analysis, histogram of $t\bar{t}$ and di-boson processes estimated by Monte Carlo simulations were subtracted from the histogram of data in Control Region, then the remaining one was considered as composed only of multi-jet events, obtaining the multi-jet shapes with a data-driven method.
Now it would be impossible to pass from an histogram-based to an event-based approach (because the network reads event-by-event information), therefore the full background estimation is directly performed.

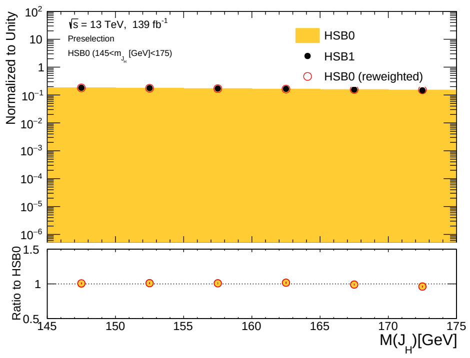
4.4.2 Re-weighted shapes in training regions

In this sections several variables before and after reweighting are shown, in the training region. m_H, $D2_{trk}$ and Anomaly Score reweighted distributions are shown

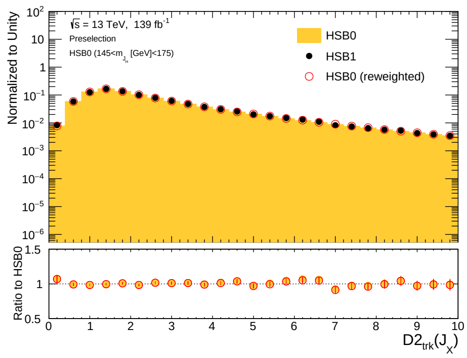
in Figures 4.23 before any region categorization. The Higgs mass distribution does not need any reweighting, therefore it is not included in following plots; D_2^{trk} and Anomaly Score are shown at this level of selection in order to understand how the variables we cut on are modeled.
In order to understand how the reweighting procedure performed, the distributions of data in HSB0 before and after the reweighting have been compared with data in HSB1 (the target distributions). This comparison has been done in the range where the network has been trained with $m_H \in (145, 175)$ GeV, separately in the Two-prong merged and resolved regions and in the Anomaly region. The bottom plots in the following figures show the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.
The reweighted shape is obtained using the nominal set of weights from the bootstrap procedure (described in section 4.6.3). Small remaining discrepancies on track jets variables exist; as they are not directly involved in the background estimation, this is not a cause of concern for final fit.

Two-prong merged

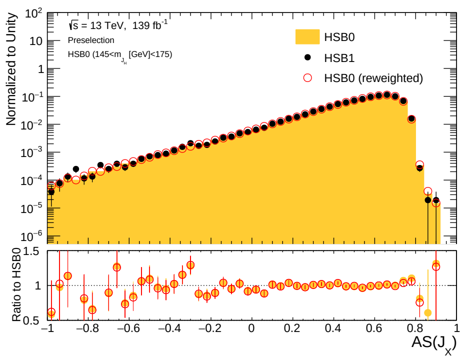
In Figure 4.24 p_T, η and E of the Higgs and X candidates large-R jets are shown, the p_T, η, and m of the first two leading trackjets associated to the Higgs candidate can be seen in Figure 4.25, while Figure 4.26 shows the reweighting effect for the X candidate mass, the final XH invariant mass and Higgs' number of charged tracks.
The bottom plots show the ratios HSB1/HSB0 before (orange) and after (red) the reweighting. The error bands in the ratios are only statistical.



(a)



(b)



(c)

Figure 4.23: m_H, X$D2_{trk}$ and X Anomaly Score in the HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

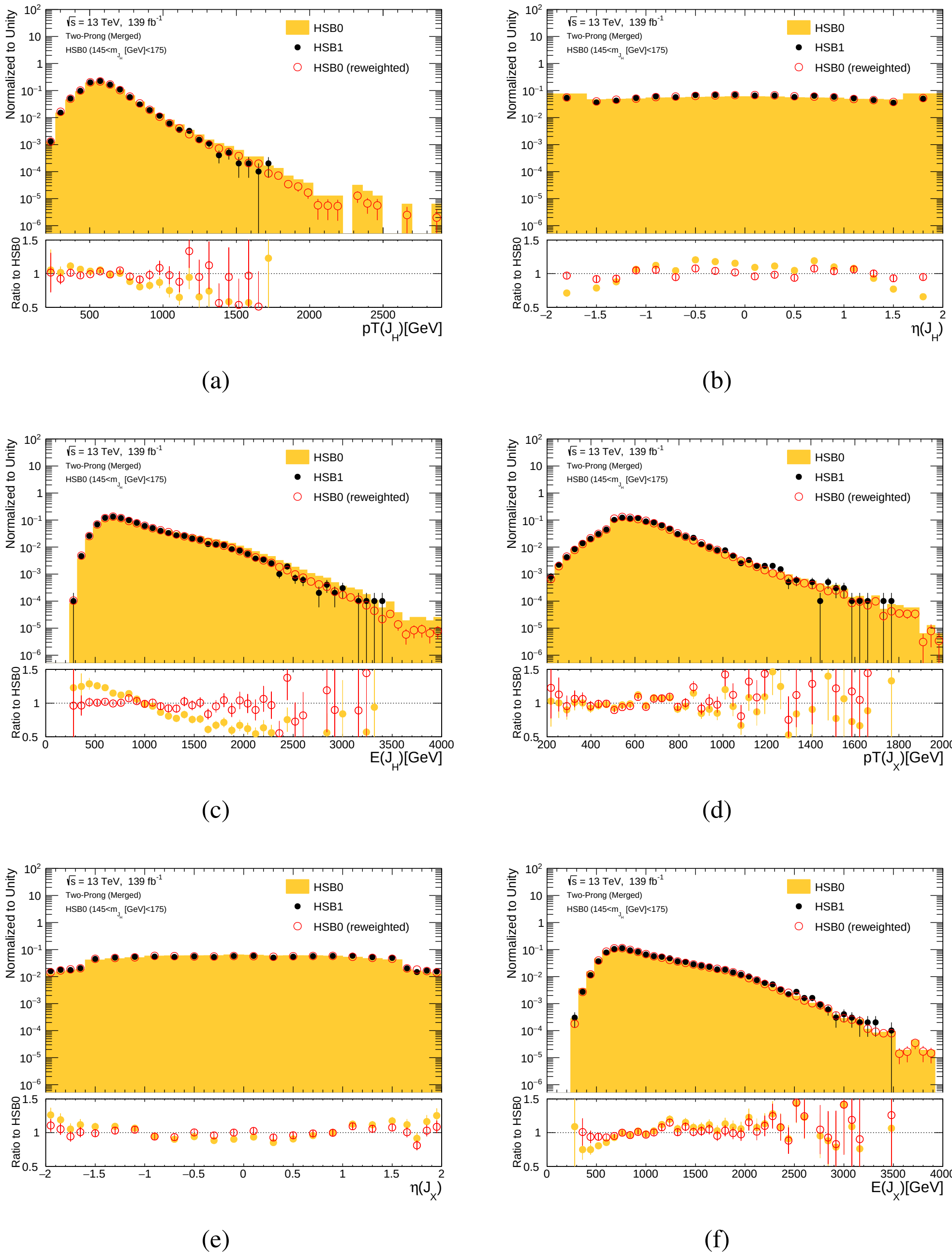
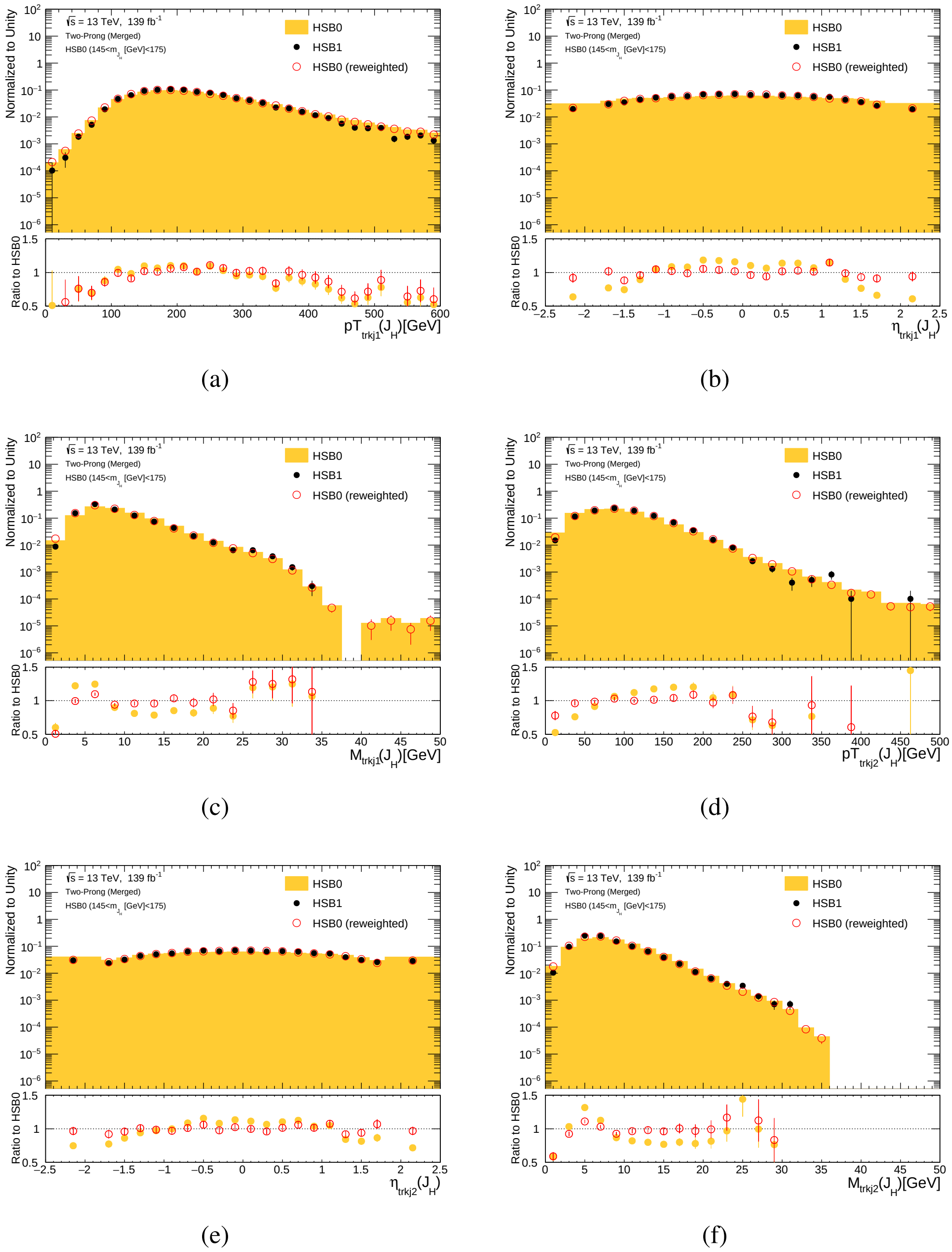
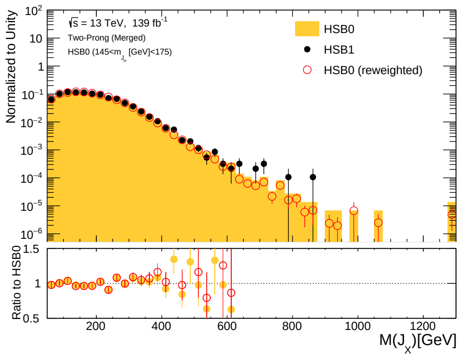


Figure 4.24: From (a) to (c) p_T, η and E of the Higgs candidate jet; from (d) to (f) p_T, η and E of the X candidate jet, in Two-prong merged HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

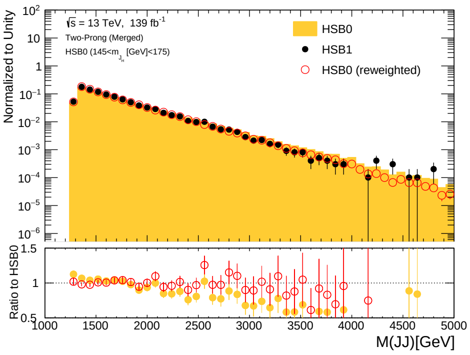


(a) (b) (c) (d) (e) (f)

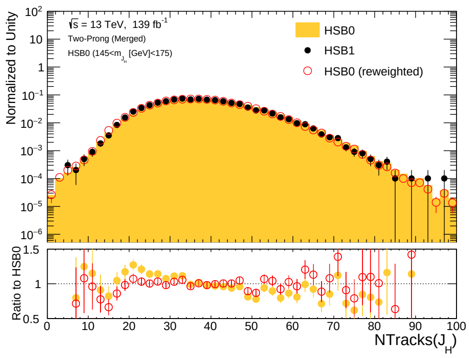
Figure 4.25: p_T, η, and mass of the first (a-c) and second (d-f) track jet (ordered in p_T) associated to the Higgs candidate fat jet, compared before and after the reweighting, in merged HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting. Small remaining discrepancies on track jets variables exist; as they are not directly involved in the background estimation, this is not a cause of concern for final fit.



(a)



(b)

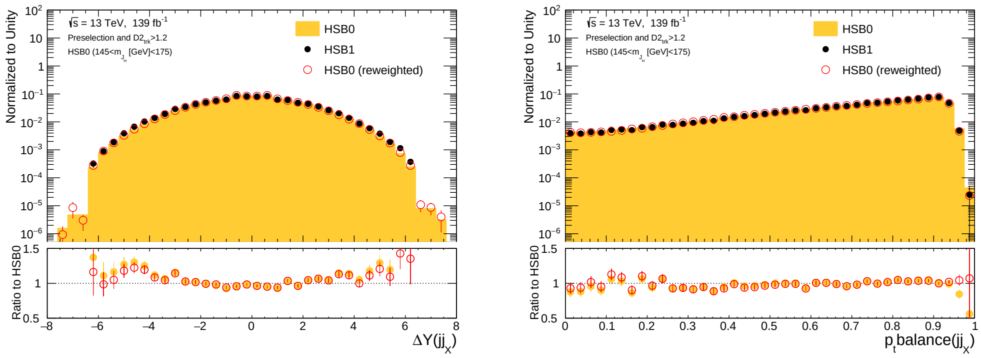


(c)

Figure 4.26: Masses of X candidate (a) and the XH dijet system in merged category (b), compared before and after the reweighting, in merged HSB. Figure (c) shows the number of charged tracks associated to the higgs candidate large-R jet. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

Two-prong resolved

Background shapes are also obtained in the resolved regime.
Small jets ΔY and p_T-balance are shown in Figure 4.27, before cutting on them to define resolved region. Mismodelling on ΔY is cutted away from $|\Delta Y| < 2.5$ request.
In Figure 4.28 there are the p_T, η, ϕ and E of the Higgs candidate large-R jet, in Figure 4.29 the p_T, η, ϕ and E of the p_T-leading and -subleading small jets associated to the X boson. The p_T, η, and m of the first two leading trackjets associated to the Higgs candidate are shown in Figure 4.30. Figure 4.31 shows the reweighting effect for the X candidate mass, the final XH invariant mass and Higgs' number of charged tracks.
The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting. The error bands in the ratios are only statistical. The re-weighted HSB0 shape is obtained using the nominal set of weights from the bootstrap procedure (described in section 4.6.3).



(a) (b)

Figure 4.27: X small jets ΔY and p_T-balance in the resolved HSB (only with $D2_{trk}(X)$¿1.2) . The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

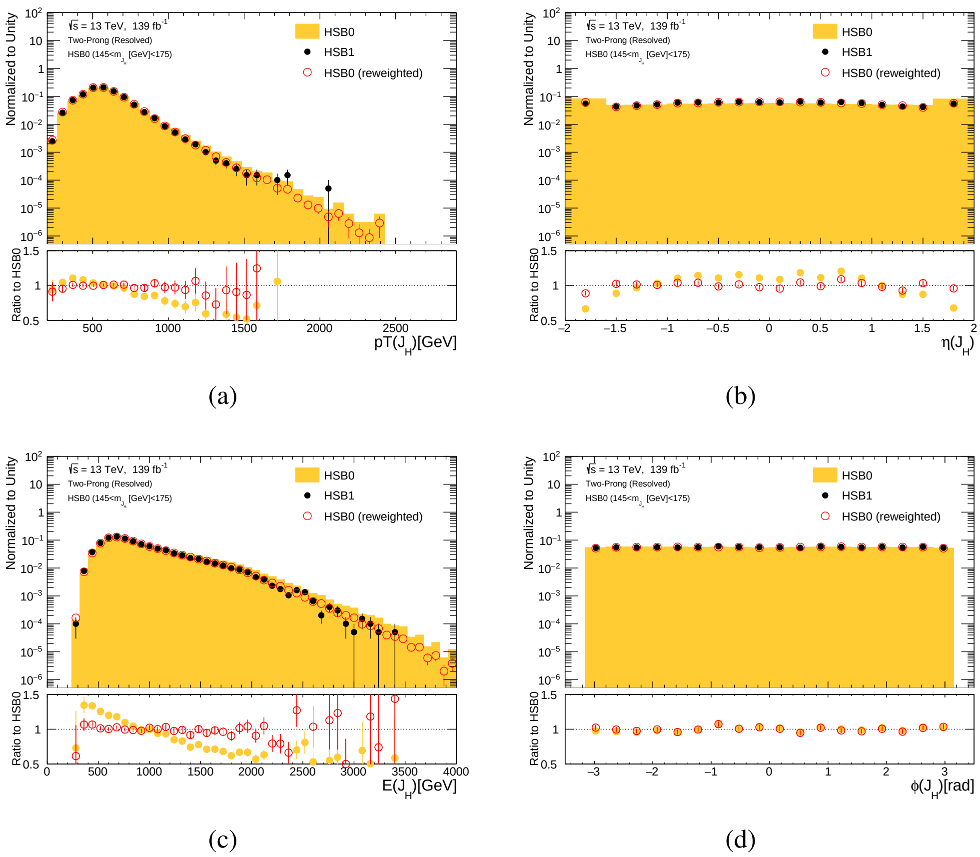


Figure 4.28: From (a) to (c) p_T, η and E of the Higgs candidate jet, in the Two-prong resolved HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

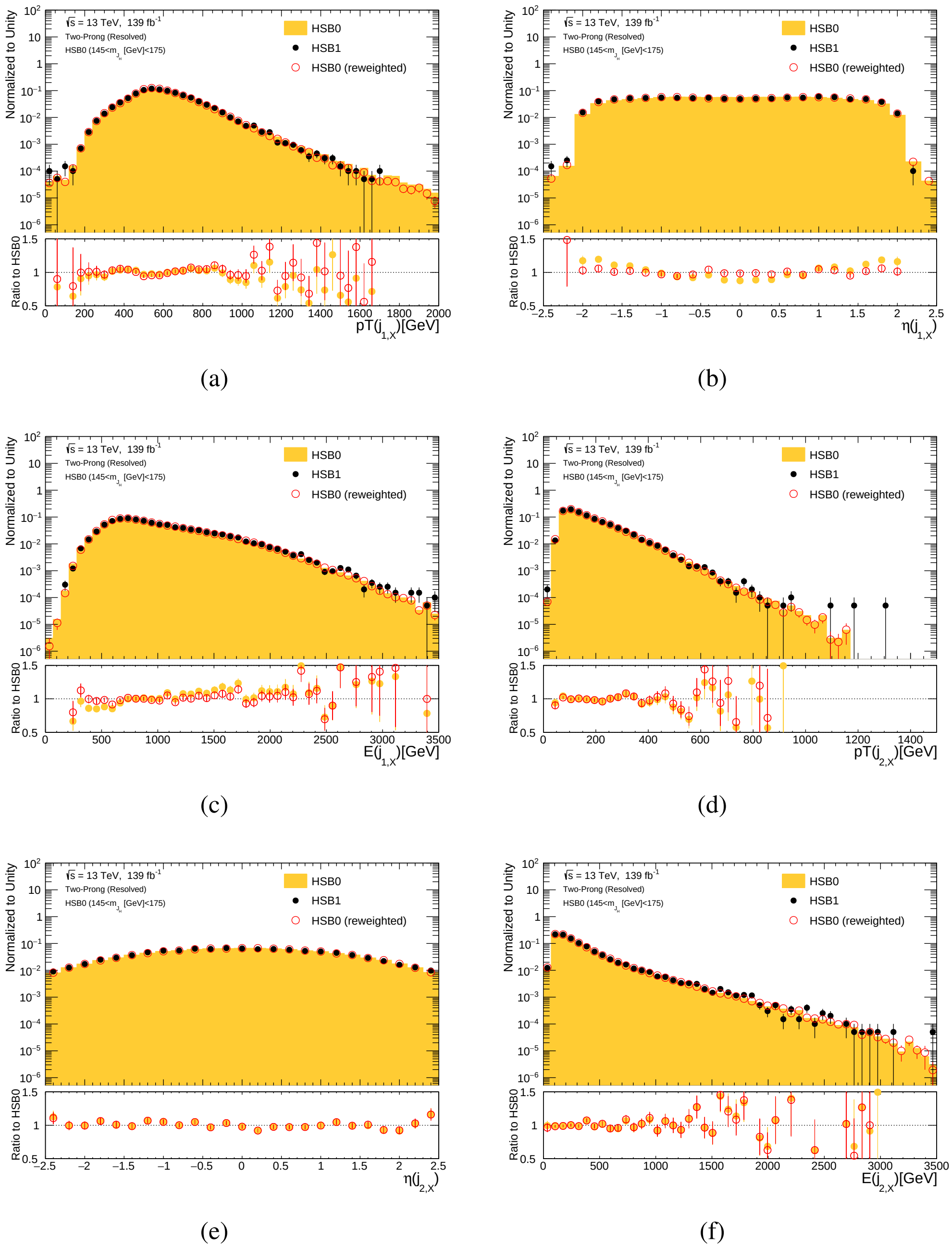
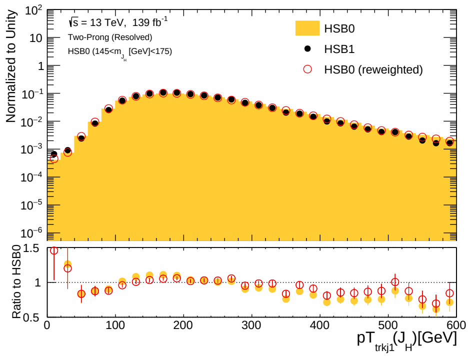
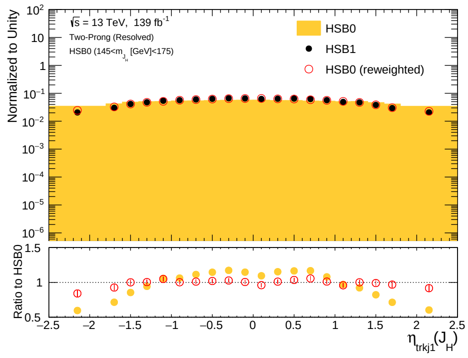


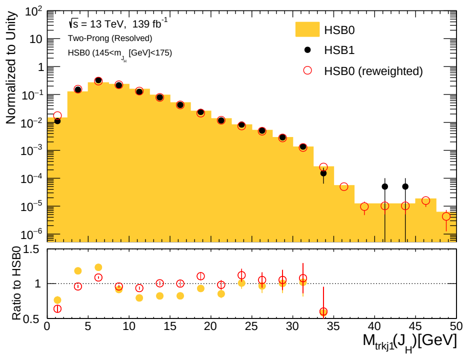
Figure 4.29: p_T (a), η (b) and E (c) of the pT-leading small jet associated to the X boson; p_T(d) η (e) and E (f) of the pT-subleading small jet associated to the X boson, in resolved HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.



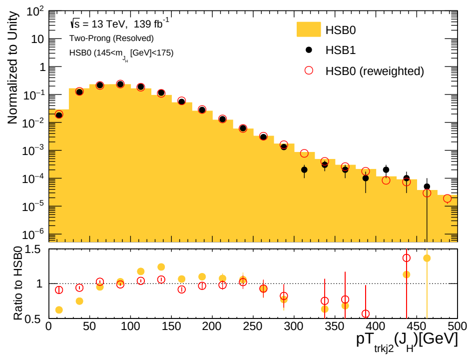
(a)



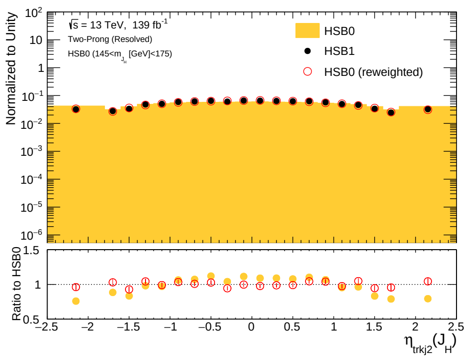
(b)



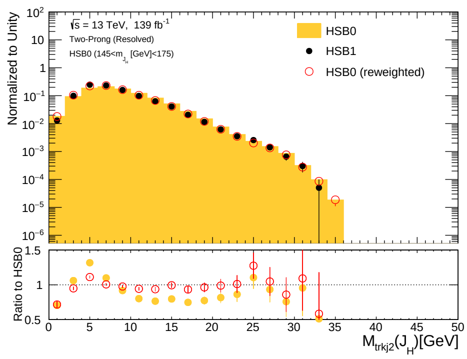
(c)



(d)

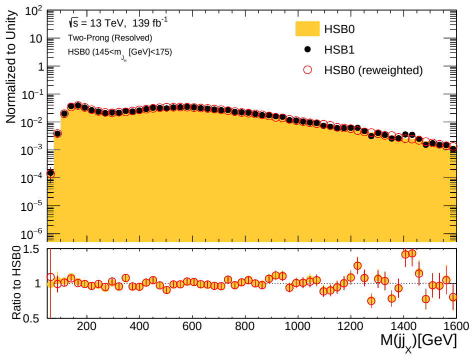


(e)

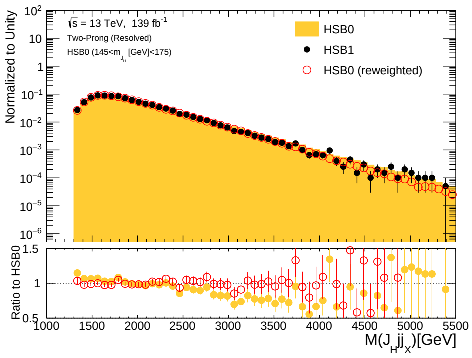


(f)

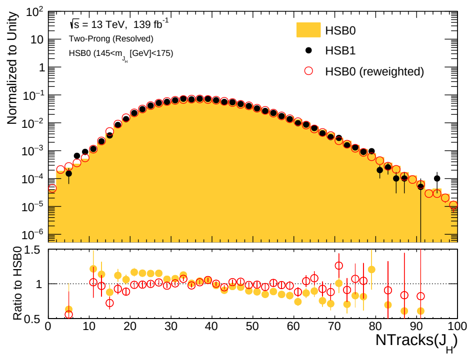
Figure 4.30: p_T, η and mass of the first (a-c) and second (d-f) track jet (ordered in p_T) associated to the Higgs candidate fat jet, compared before and after the reweighting, in resolved HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting. Small remaining discrepancies on track jets variables exist; as they are not directly involved in the background estimation, this is not a cause of concern for final fit.



(a)



(b)

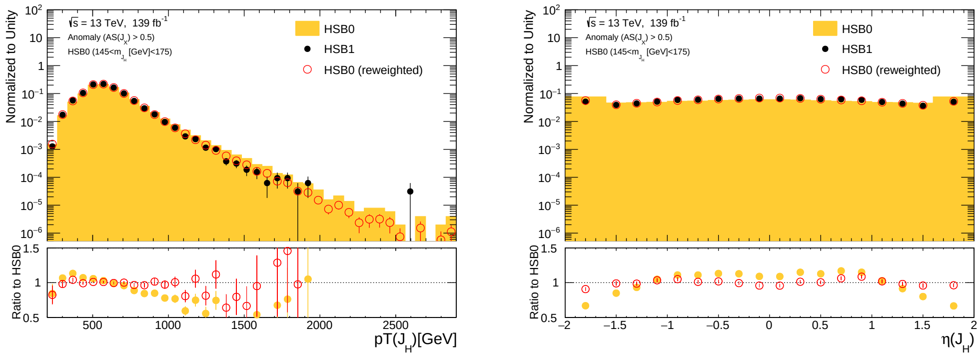


(c)

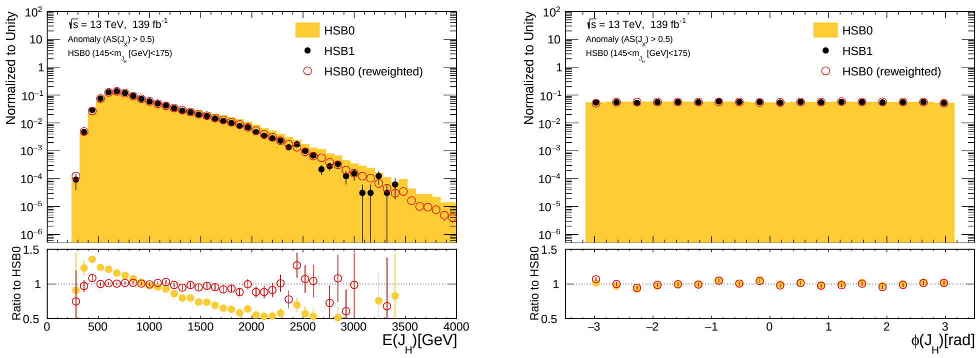
Figure 4.31: Masses of X candidate reconstructed with the two small jets(a) and the XH dijet system in resolved category (b), compared before and after the reweighting, in resolved HSB. Figure (c) shows the number of charged tracks associated to the higgs candidate large-R jet. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

Anomaly region

Comparisons of kinematic distributions have been done in the Anomaly region as well. Figure 4.32 and 4.33 show respectively p_T, η, ϕ and E of the Higgs and X candidate large-R jets along with m_{J_X} and m_{JJ}.
The bottom plots show the ratios HSB1/HSB0 before (orange) and after (red) the reweighting. The error bands in the ratios are only statistical.



(a) (b)



(c) (d)

Figure 4.32: p_T (a), η (b), E (c) and ϕ (d) of the Higgs candidate fat jet, compared before and after the reweighting, in anomaly HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

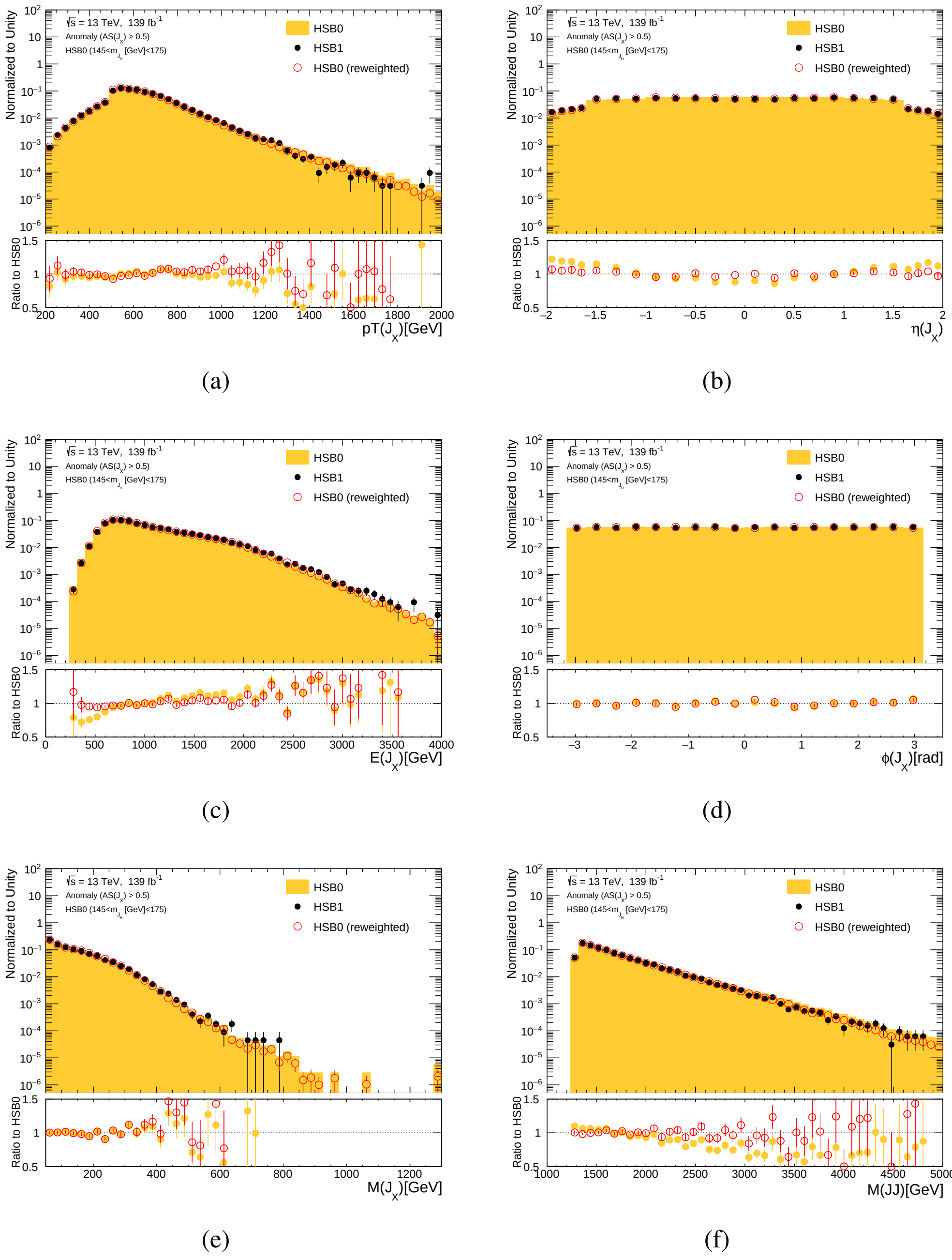


Figure 4.33: p_T (a), η (b), E (c) and ϕ (d) of the X candidate fat jet; masses of X candidate (e) and the XH dijet system (f) in Anomaly HSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

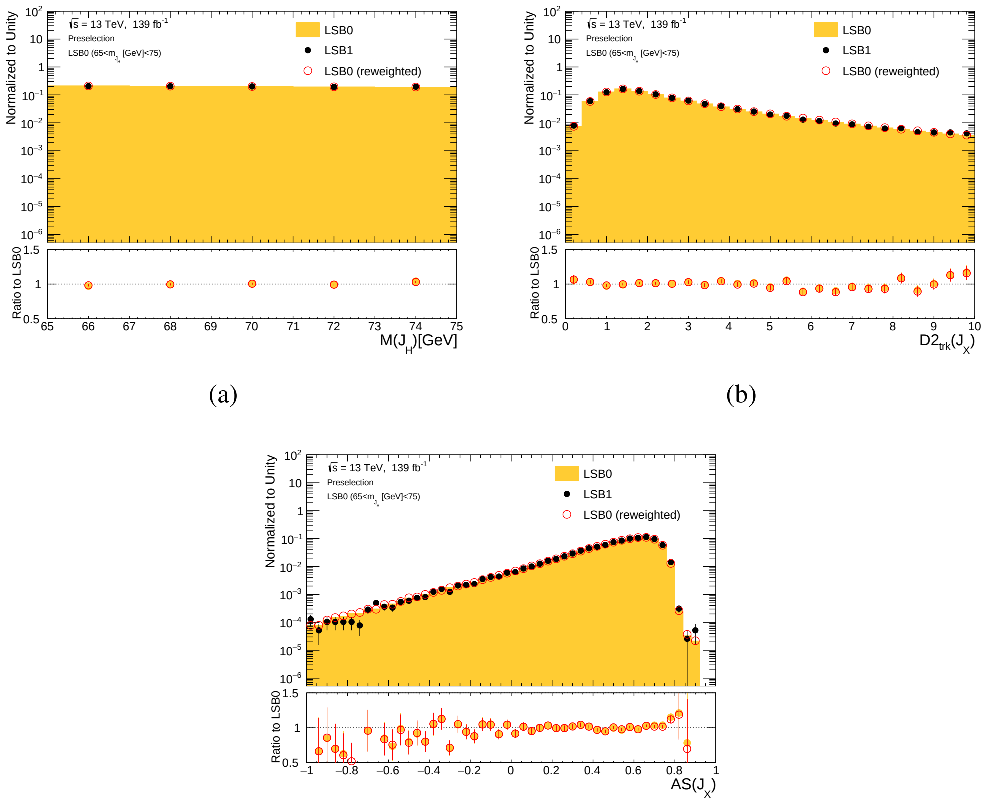
4.4.3 Validation Region (LSB)

The validation of the NN-derived weights for the background estimation has been performed in the low sideband (LSB) of the Higgs candidate jet mass. Actually not the full LSB but only the range $65 - 75$ GeV has been used, since in the kinematic region below 65 GeV the Higgs candidate fat jet reconstruction suffers for the proximity to the lower limit cut on the jet mass at 50 GeV, therefore the assumption of the independence of ratio between tagged and untagged regions from the Higgs mass is no longer satisfied.
m_H, $D2_{trk}$ and Anomaly Score reweighted shapes are shown in Figures 4.34 before any region categorization. The Higgs mass distribution does not need any reweighting, therefore it is not included in following plots; $D2_{trk}$ and Anomaly Score are shown at this level of selection in order to understand how the variables we cut on are modeled.

Two-prong merged

In this subsection normalized plots for several variables are shown, comparing the distributions of data among LSB1, LSB0 and re-weighted LSB0 (which is the estimated shape of LSB0), with $D2_{trk} < 1.2$ cut to define the merged category. Figures 4.35 to 4.37 show closure of background prediction to data for the same variables shown in Section 4.4.2.
The error bands in the ratios are only statistical. The re-weighted LSB0 shape is obtained using the nominal set of weights from the bootstrap procedure (described in section 4.6.3).
Again, generally good closure is observed of the background prediction to data. There is a remaining discrepancy in track-jets p_T(in particular at low values of the leading jet) and on track jets mass in LSB due to the fact that the assumption of independence from m_H is not perfectly respected for track-jets variables (as shown at the beginning of this Section). Since the discrepancy grows as the m_H assumes lower values, from the right to the left on the m_H axis, it is reasonable to suppose track-jets variables are well reweighted in Higgs mass windows as they are in HSB.

Non-closure of the data-driven background estimation is taken as a systematic uncertainty of the background estimation method, as detailed in Section 4.6.3. The remaining discrepancy on track jets variables are not a cause of concern for final fit.

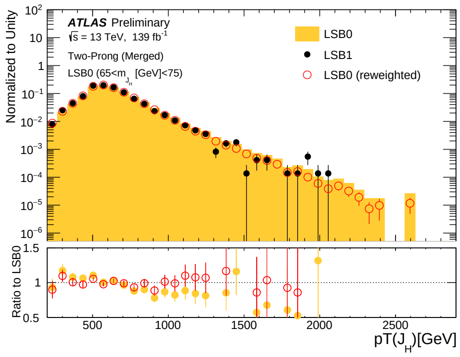


(a)

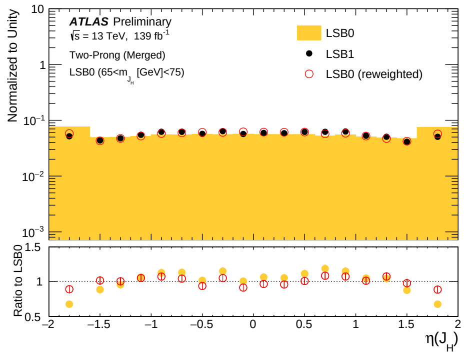
(b)

(c)

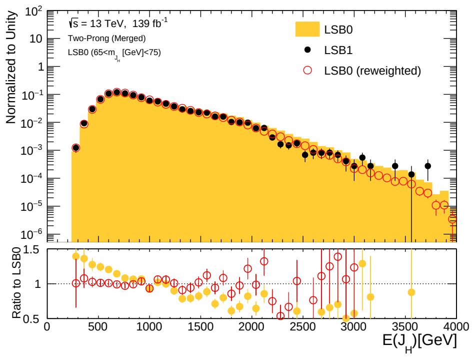
Figure 4.34: m_H, X $D2_{trk}$ and X Anomaly Score in the LSB. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.



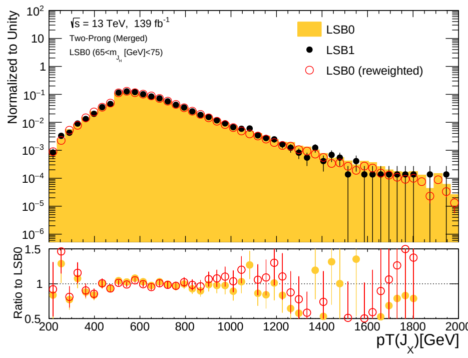
(a)



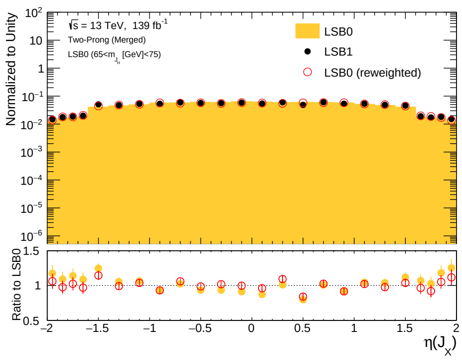
(b)



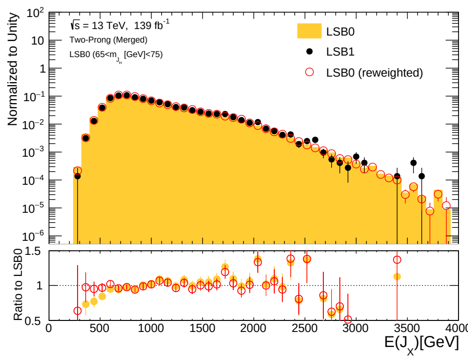
(c)



(d)



(e)



(f)

Figure 4.35: p_T (a), η [7] (b) and E (c) of the Higgs candidate and (from (d) to (f)) of the X candidate large-R jet in merged LSB. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.

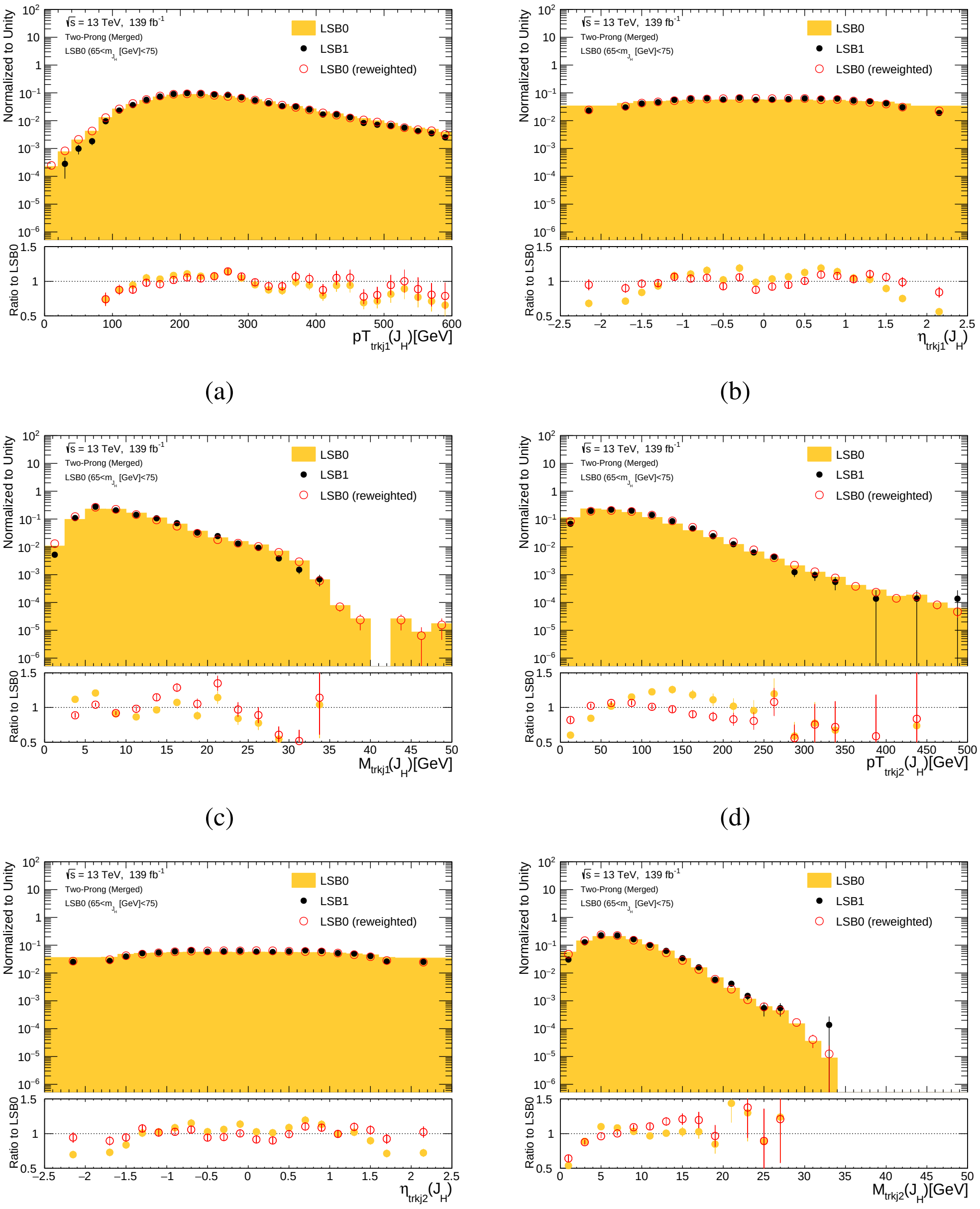
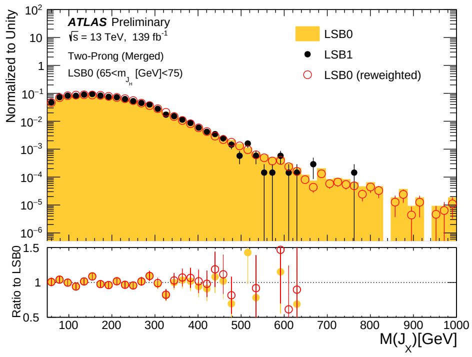
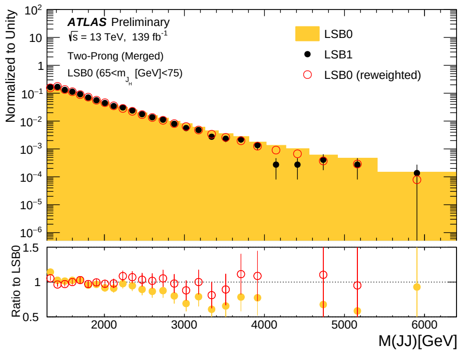


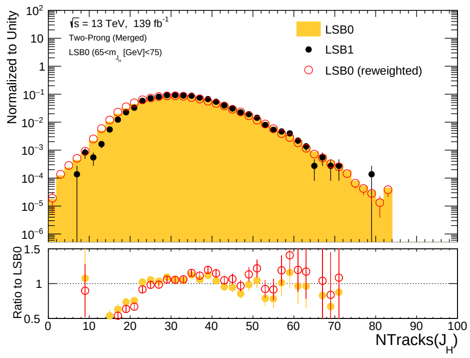
Figure 4.36: p_T, η, and mass of the first two track jets (ordered in p_T) associated to the Higgs candidate fat jet, compared before and after the reweighting, in merged LSB. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.



(a)



(b)



(c)

Figure 4.37: Masses of X candidate (a) and the XH dijet system in merged category (b) [7], compared before and after the reweighting, in merged LSB. Figure (c) shows the number of charged tracks associated to the higgs candidate large-R jet. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.

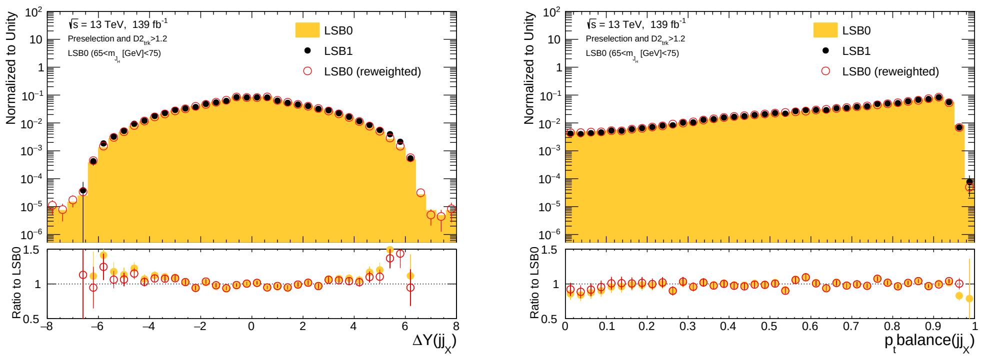
Two-prong resolved

Small jets ΔY and p_T-balance are shown in Figure 4.38, before cutting on them to define resolved region. Mismodelling on ΔY is cutted away from $|\Delta Y| < 2.5$ request.

Compared distributions of data among LSB1, LSB0 and re-weighted LSB0 (which is the estimated shape of LSB0), with $D2_{trk} > 1.2$ cut to define the resolved category, are shown in this subsection. Figures 4.39 to 4.42 show closure of background prediction to data.

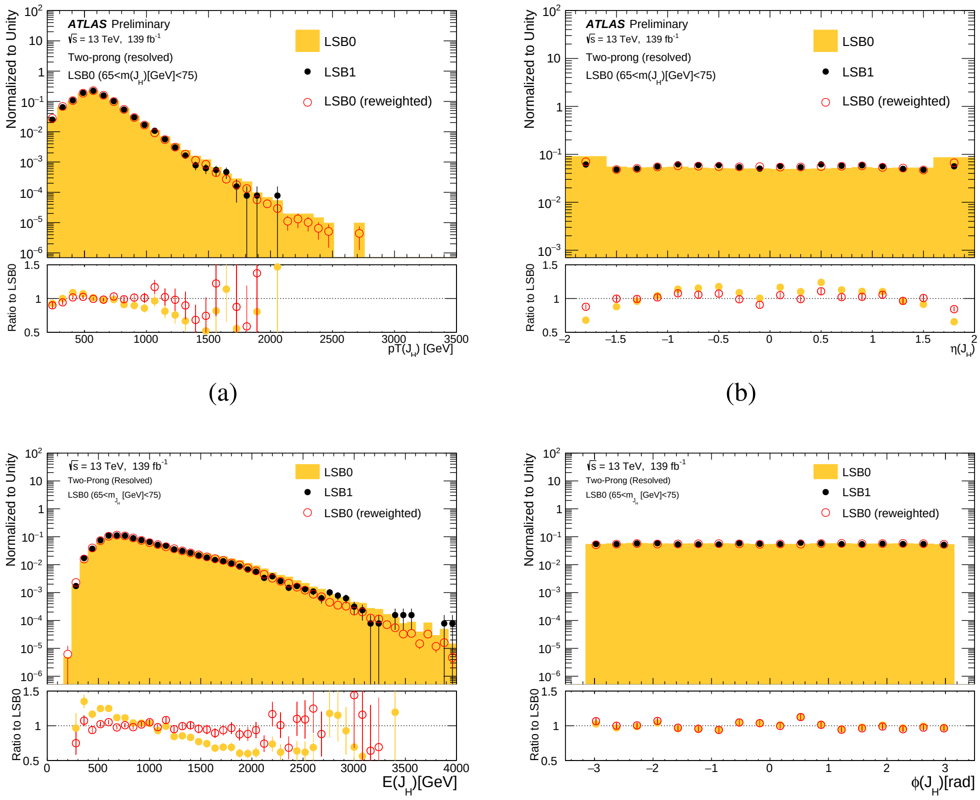
The error bands in the ratios are only statistical. The re-weighted LSB0 shape is obtained using the nominal set of weights from the boostrap procedure (described in section 4.6.3).

The same observations done in merged LSB can be applied here.



(a) (b)

Figure 4.38: X small jets ΔY and p_T-balance in the resolved LSB (only with $D2_{trk}$¿1.2) . The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.



(a) (b)

(c) (d)

Figure 4.39: p_T (a), η(b), E (c) and ϕ (d) of the Higgs candidate fat jet, compared before and after the reweighting, in resolved LSB. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.

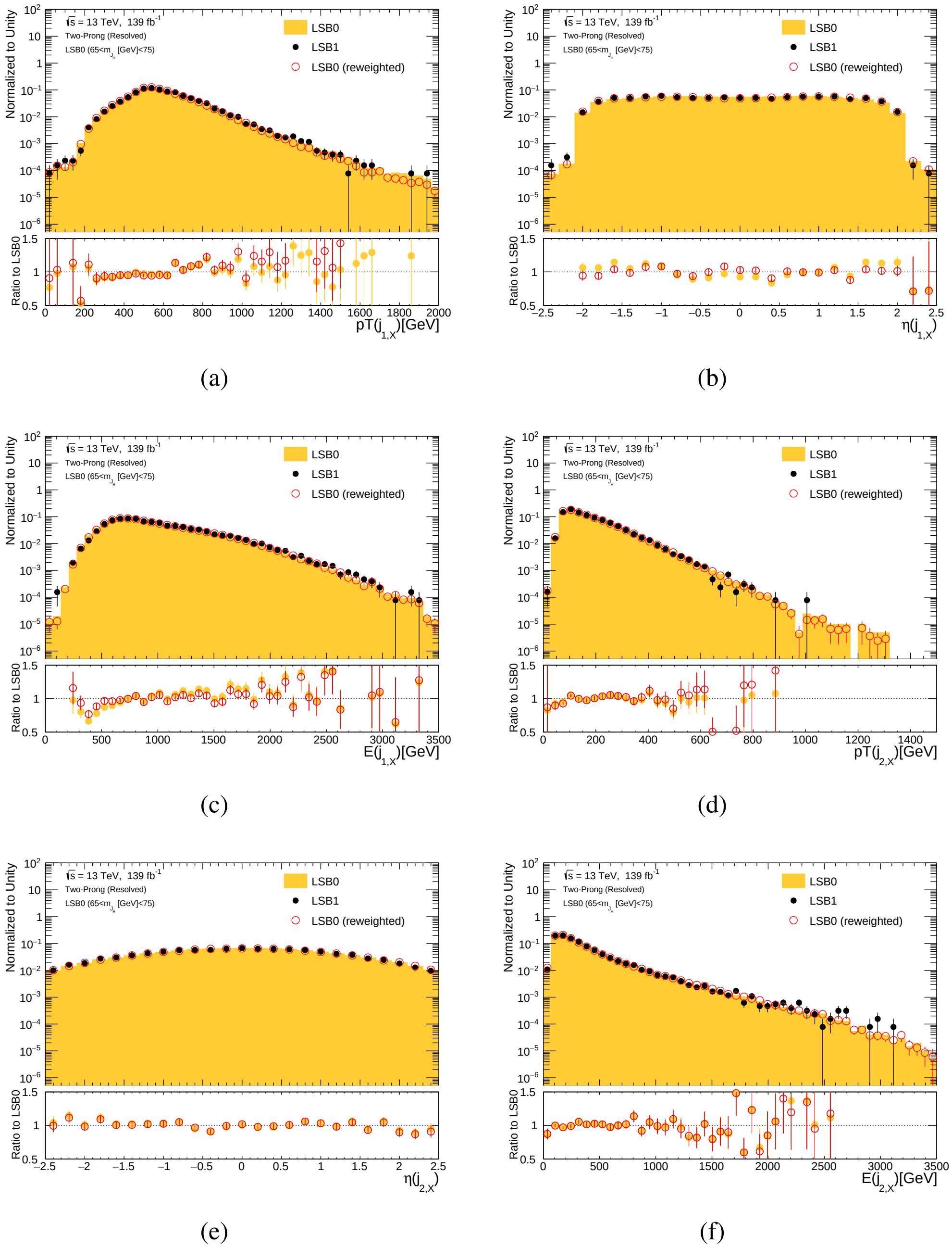


Figure 4.40: p_T (a), η (b) and E (c) of the p_T-leading small jet (from (d) to (f) of the p_T-subleading) associated to the X boson, in resolved LSB. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.

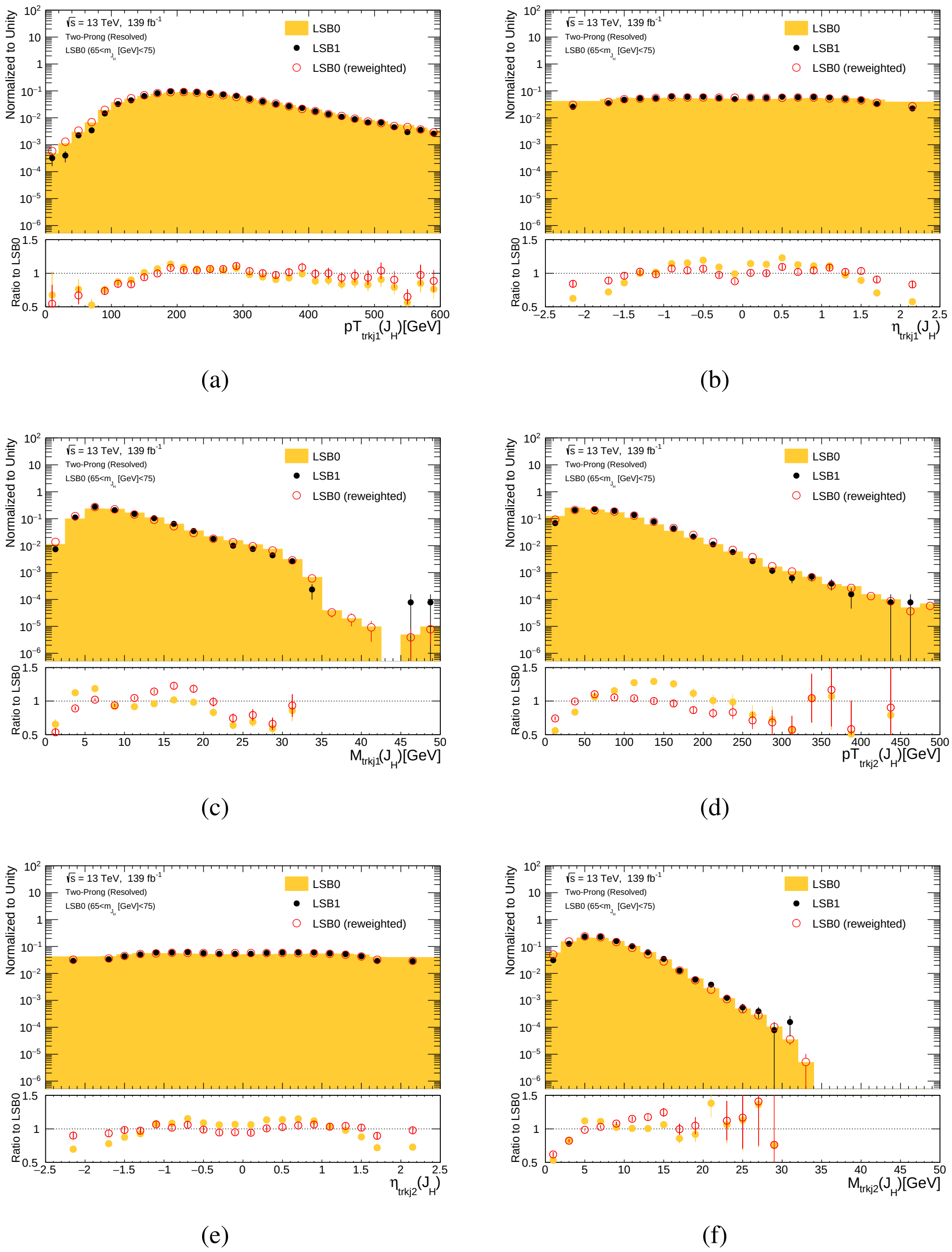
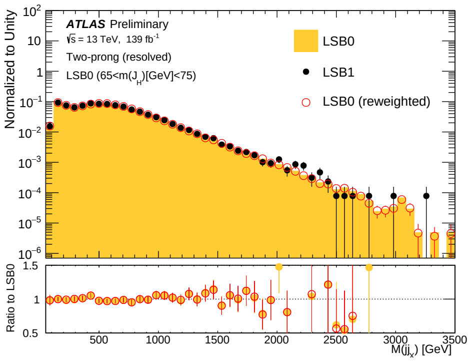
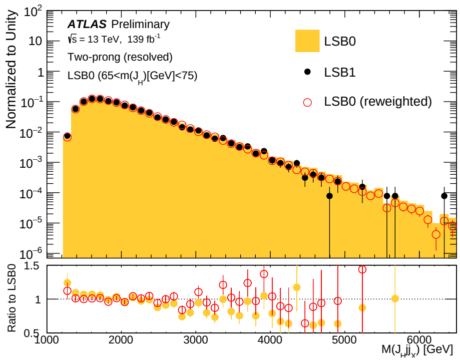


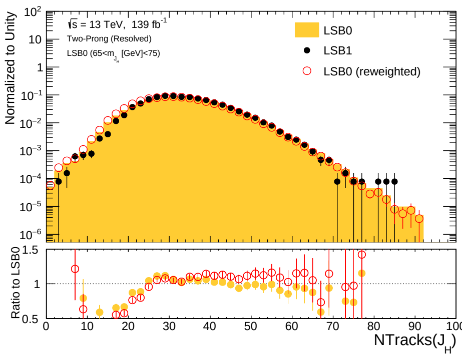
Figure 4.41: p_T, η, and mass of the first two track jets (ordered in p_T) associated to the Higgs candidate fat jet, compared before and after the reweighting, in resolved LSB. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.



(a)



(b)

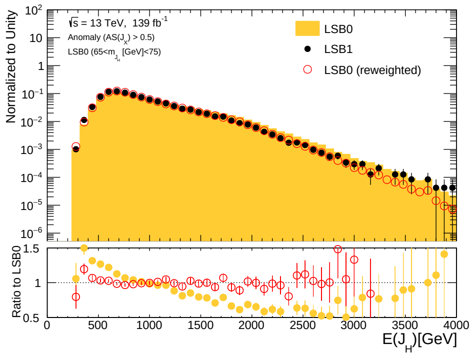


(c)

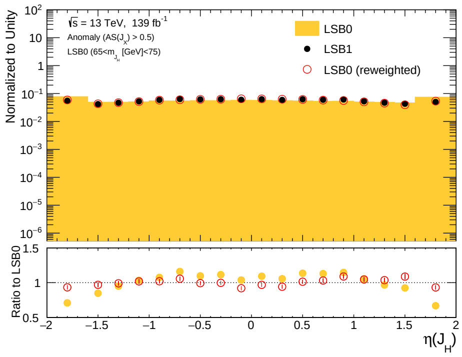
Figure 4.42: Masses of X candidate reconstructed with the two small jets(a) and the XH dijet system in resolved category (b), compared before and after the reweighting, in resolved LSB. Figure (c) shows the number of charged tracks associated to the higgs candidate large-R jet. The bottom plots shows the ratios LSB1/LSB0 before (orange) and after (red) the reweighting.

Anomaly Score Region

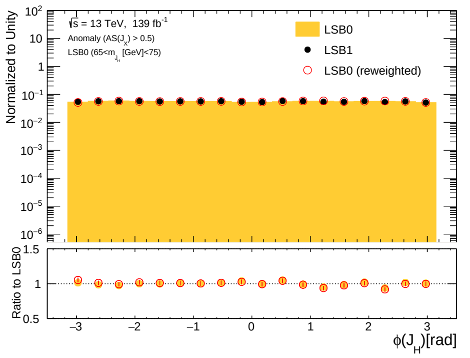
The AS region background estimation is validated to data in the LSB as shown in Figures 4.43 and 4.44, showing variables related to the Higgs kinematics, X kinematics, and m_{JJ}/X anomaly score/X mass, respectively. No significant variations are observed.



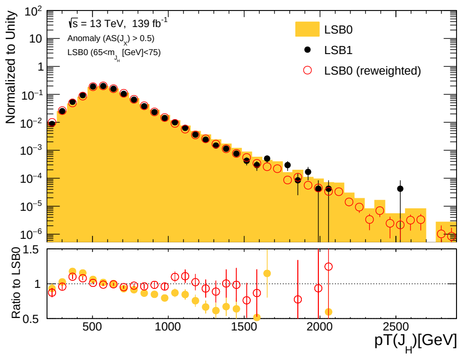
(a)



(b)



(c)



(d)

Figure 4.43: p_T (a), η(b), E (c) and ϕ (d) of the Higgs candidate fat jet, compared before and after the reweighting, in anomaly LSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

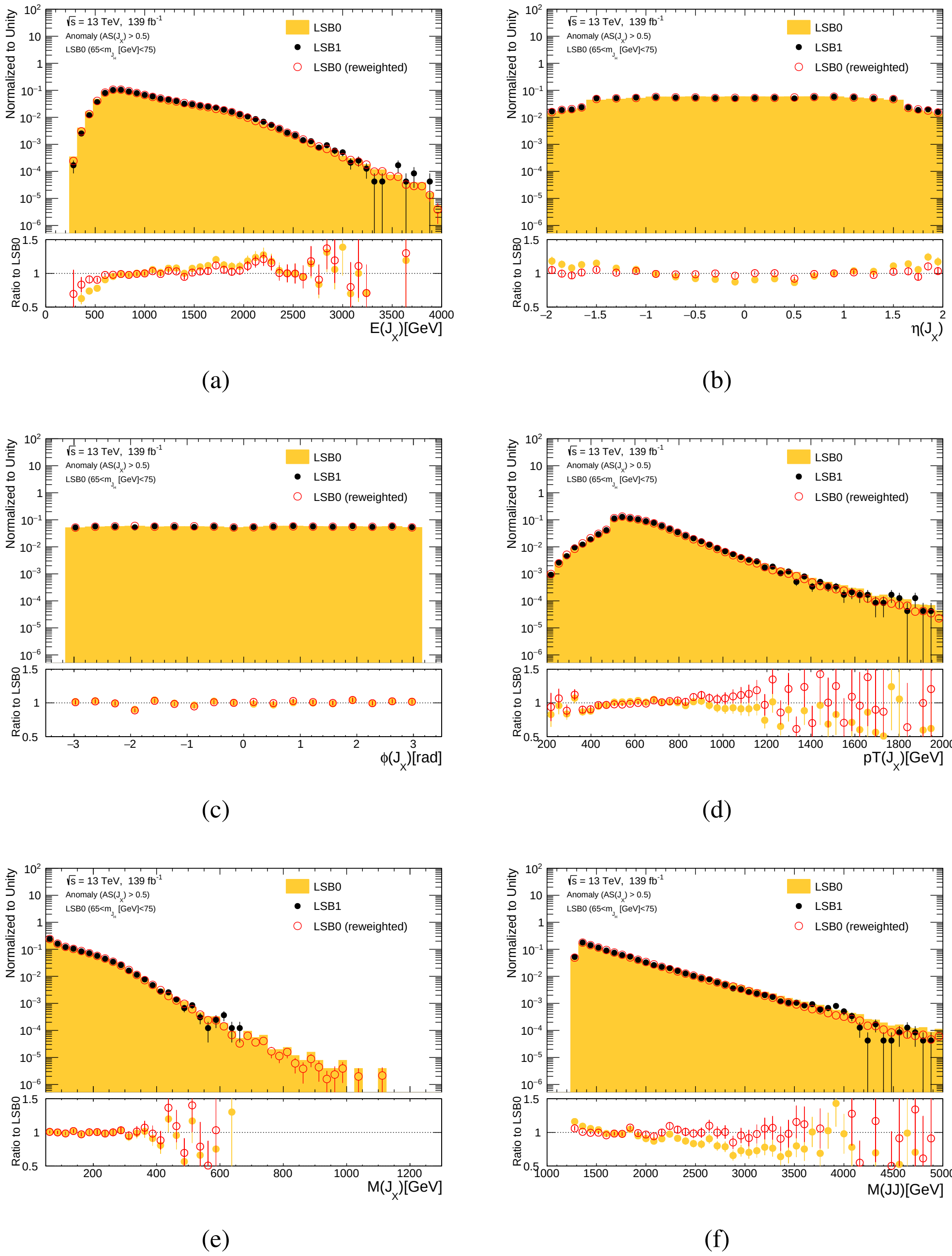


Figure 4.44: p_T (a), η(b), E (c) and ϕ (d) of the X candidate fat jet; masses of X candidate (e) and the XH dijet system (f), in Anomaly LSB. The bottom plots shows the ratios HSB1/HSB0 before (orange) and after (red) the reweighting.

4.5 Signal mass resolution & interpolation

In order to optimize the binning of the fit variables m_{JJ} and m_X, the mass resolution for both Y and X particles is studied using double sided Crystal Ball functions. Each generated sample mass is fitted and the resolution is defined to be the width of the core. As the resolution is expected to degrade with increasing mass, this study motivates the choice of binning in Y and X mass for the statistical procedure. Figure 4.45 shows the obtained resolution for every signal point, with a linear fit as a function of both X and Y masses. The final binning in the analysis is chosen

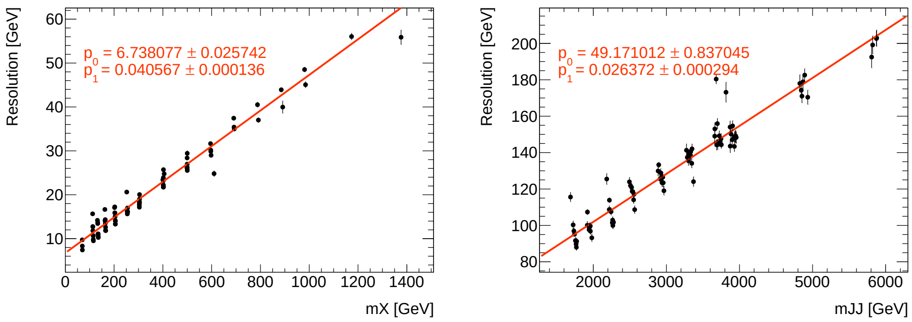


Figure 4.45: A linear fit of signal point mass resolution to the mass of the X (left) and the mass of the Y (right).

such that an excess in m_{JJ} or m_X will be detectable over the smoothly falling background, as described in Section 4.7.

To provide results that are more finely-grained in m_{JJ} and m_X than is possible through full simulated signal points, an interpolation method has been used. This method takes the existing simulated mass distributions, parameterizes their shape using a unique probability distribution function (PDF) and uses the PDF to predict the signal PDF of intermediate mass points using non-linear moment morphing techniques [40].
The mass distributions are fitted with non-custom functions by a kernel density estimation (KDE) [72]. These were found to produce more accurate sample modeling than Crystal Ball functional fits, which are only used to determine the binning and these parameters are not involved in the morphing. Figure 4.46 gives examples of the resulting KDE PDF for two $Y \to XH$ signal points.
Interpolation is then done by moment morphing technique. The m_{JJ} distribution that is fit must be morphed in two dimensions. First, real MC signal histograms are morphed across bins of m_X. The morphed shapes are then used as inputs (along with real MC signal) to a second morphing procedure that builds intermediate

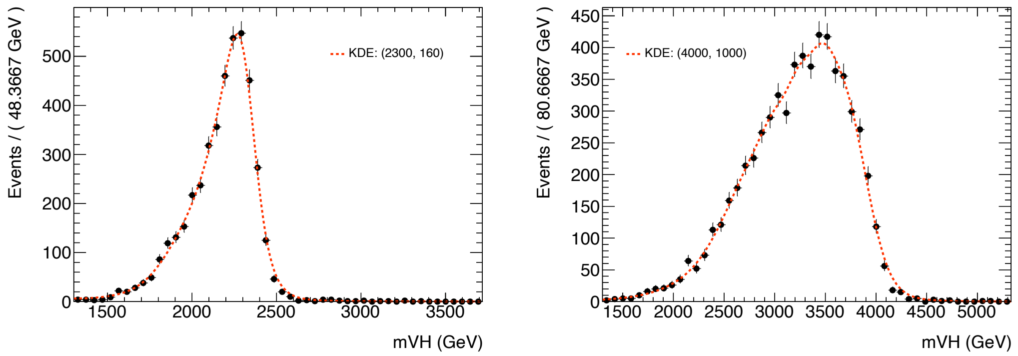


Figure 4.46: Example KDE fits to m_{JJ} shapes in real MC signal simulation, for sample points (m_Y,m_X) = (2300, 160) GeV (left) and (4000, 1000) GeV (right).

histograms in bins of m_{JJ}.
Figure 4.47 shows the resulting intermediate m_{JJ} shapes for both dimensions of the morphing procedure, given two boundary histograms. The density of morphed signals is determined by the interpolation parameter α, which has units of GeV.

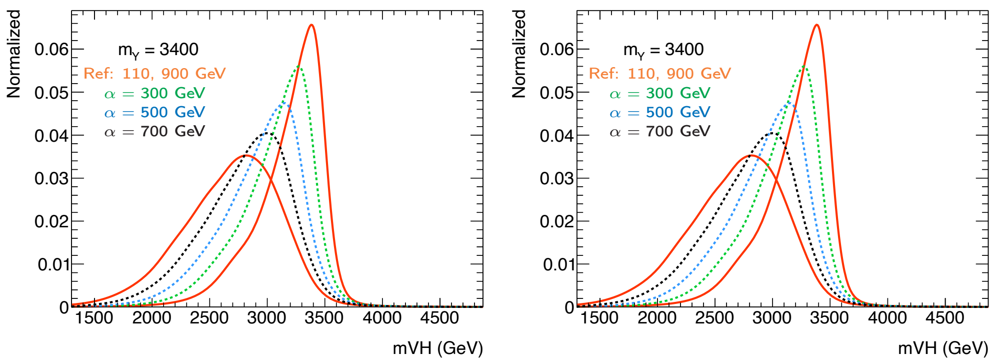


Figure 4.47: Examples of morphing between two real signal shapes (red) to a variety of intermediate morphing paramter values α [GeV], for the first step in morphing across bins of m_X (left) and the second step across bins of m_{JJ} (right).

Once the shapes of the morphed histograms are obtained, their normalization needs to be determined by a yield interpolation procedure. The simulated signal yields are nonlinear with respect to m_Y, and need to be modeled by piecewise cubic functions. Figure 4.48 shows the result of this functional fitting for an example (m_Y, m_X) window of (3000, 500) GeV. The observed trend across m_Y and m_X is attributed to the inefficiency of the merged selection for more resolved points. This can be seen more clearly in the two-dimensional yields shown in Figure 4.19. The normalization for the morphed m_{JJ} histograms is taken from these functions for the desired m_Y and m_X values.

Figure 4.49 shows examples of the closure between the morphing procedure and

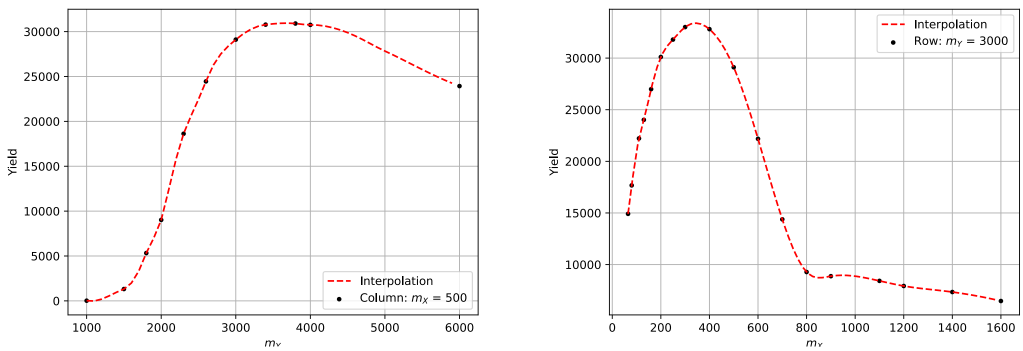


Figure 4.48: MC signal yields for an example grid bin of (m_Y, m_X) = (3000, 500) GeV, along with the piecewise cubic function used to determine the necessary normalization for the interpolated signal points.

the real MC signal simulation. No significant deviations or trends in the ratio are present, indicating good agreement in both normalization and shape. These morphed m_{JJ} histograms are used directly as inputs to the statistical framework to determine the finely-grained exclusion result.

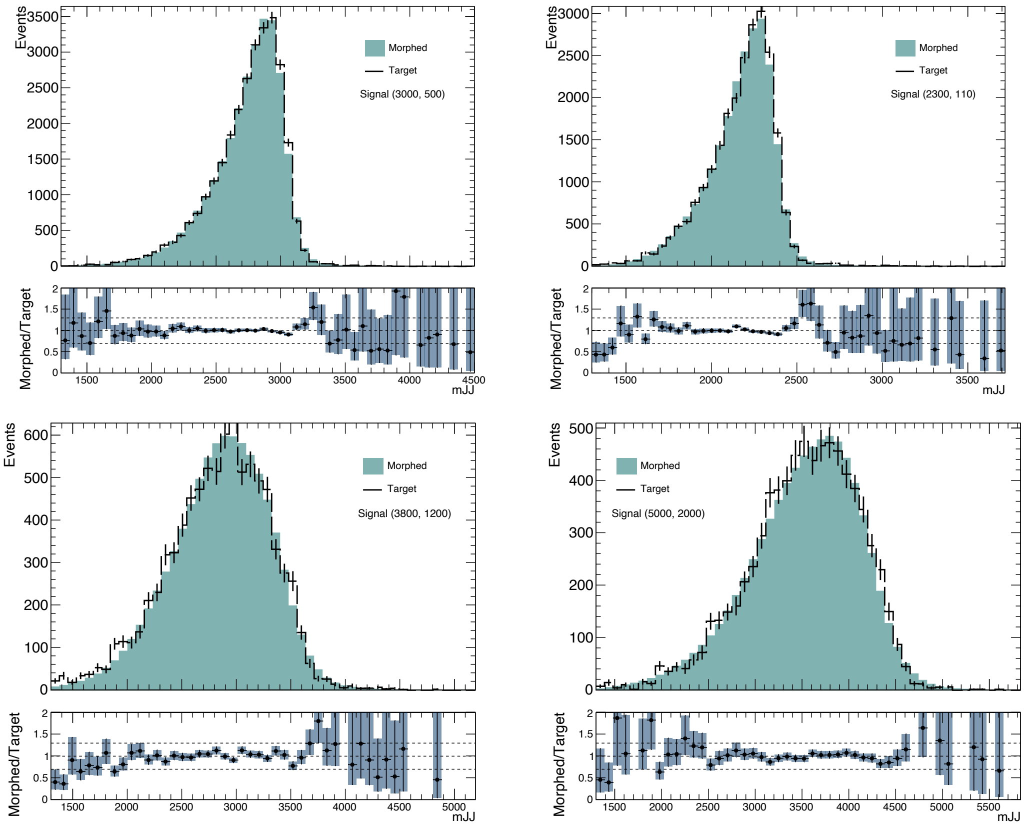


Figure 4.49: Comparisons of the morphed m_{JJ} shape to the real signal MC distribution, for four example $Y \to XH$ signal points: (m_Y, m_X) = (3000, 500) (top-left), (2300, 110) (top-right), (3800, 1200) (bottom-left), and (5000, 2000) (bottom-right).

4.6 Systematics uncertainties

In this section the systematics uncertainties considered in the analysis are discussed, starting from the theoretical one, moving to those related to the signals and then to the data-driven background.

4.6.1 Theory uncertainties

Two sources of theory systematic uncertainties are considered: from the parton distribution functions and from the Monte Carlo simulations tunes.

PDF uncertainties Uncertainties in the behavior of the parton distribution functions at the high Q^2 values explored in this analysis can potentially have a large effect on signal production rates. This systematic uncertainty is derived by applying the methodology outlined by the PDF4LHC group using event level reweighting to three additional PDF sets: CT14, MMHT2014, and NNPDF3.0.
The PDF variations here can be either a comparison to different PDF sets, as well as comparisons to variations in the error sets of the nominal PDF.

ISR/FSR uncertainties Systematic uncertainties are inherent in the choice of Monte Carlo tune. Five pairs of systematic variations are used to cover uncertainties in the A14 tuning parameters describing Initial State Radiation (ISR), Final State Radiation (FSR), and Multi-Parton Interaction (MPI). These are implemented via Pythia variations imposed at the evgen level.
The uncertainty on the signal acceptance is evaluated at truth level, before boson tagging cuts, and found to be both small and consistent across signal phase space. A conservative flat 3% is chosen across the signal grid.

4.6.2 Signal/MC uncertainties

Uncertainties applied to the signal relate to experimental features, such as luminosity or combined performance/reconstruction, or in the modeling of the simulated signal m_{JJ} shape in the fit.
Signal uncertainties are treated as correlated across the merged and resolved signal regions.

The uncertainty from the trigger selection is negligible, as the requirement on $m_{JJ}>$ 1.3 TeV ensures that the trigger is fully efficient, and thus not included.

Luminosity

A flat 1.7% uncertainty on signal MC has been applied [34], as recommended and measured with LUCID [39].

Jet uncertainties

Jet uncertainties are implemented using standard variations produced by the ATLAS jet/E_t^{miss} combined performance (CP) group.
The size of the total jet p_T resolution (JpTR) is taken to be 2%.

Uncertainty in the jet $p_T$scale (JpTS) is an important effect in the search for resonant structures in the presence of rapidly falling background spectra. This uncertainty shifts the expected signal mass spectrum, particularly the peak of the resonance, affecting the significance of an excess if observed. The JpTS uncertainty is evaluated using track-to-calo double ratios between the data and the MC, as the ratio of two measures of jet p_T are expected to be the same in both data and MC, so any observed differences are assigned as baseline systematic uncertainties. Additional uncertainties due to the reconstruction and modelling of tracks are taken into account as well. The following list of mass and $p_T$scale uncertainties for TCC trimmed jets are implemented:

- Baseline
- Modelling
- Tracking
- TotalStat
- Closure

Example plots of these variation shapes on m_{JJ} for the (m_Y=3400, m_X=500) GeV signal point in the boosted SR can be seen in Figure 4.50 and 4.51 for the p_T and mass variations, respectively.
The size of the total JpTS uncertainty varies with p_T and η and is typically around $\pm$ 50%.
Even though the analysis relies on jets build from TCC objects, the total $p_T$of the jet is still solely derived from calorimeter information, keeping it independent of the track-based jet p_T. Past analyses have studied the possible impact of calorimeter vs. track-based p_T by cross-calibrating between per-jet TCC and LCTopo p_T and found it to be negligible [35].
Small-R jet scale and resolution uncertainties are similarly estimated through data to MC comparisons and in-situ corrections [63].

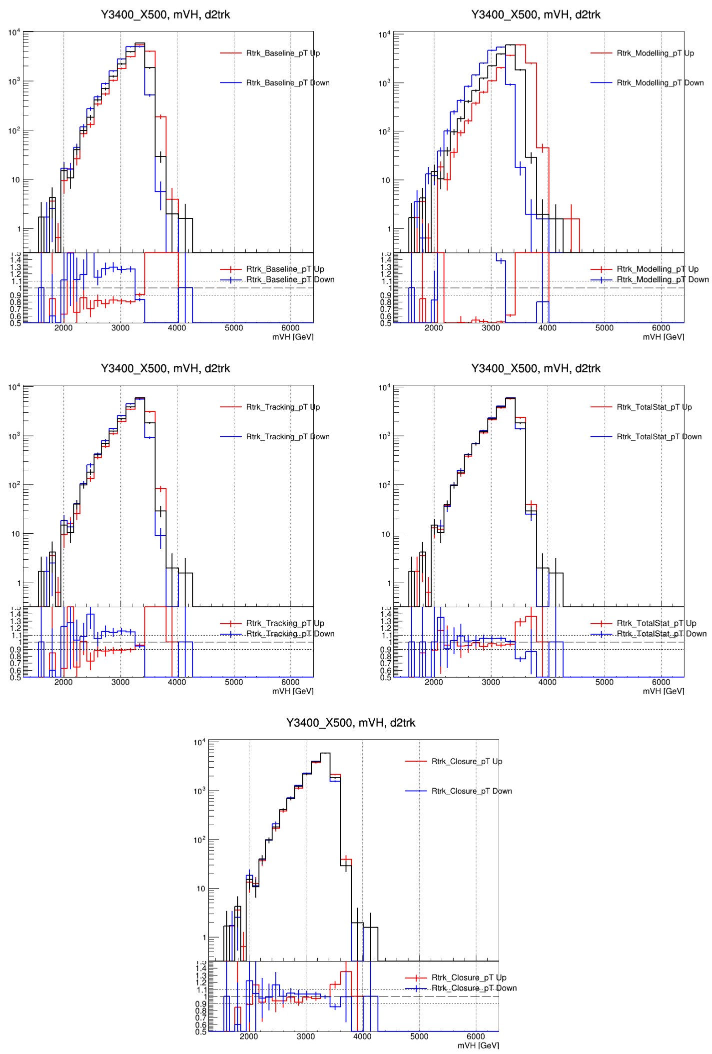


Figure 4.50: m_{JJ} distributions of the five RTrack jet CP variations (in order: Baseline, Modeling, Tracking, TotalStat, Closure) for an example $Y \to XH$ signal point of (m_Y=3400, m_X=500) GeV, on p_T.

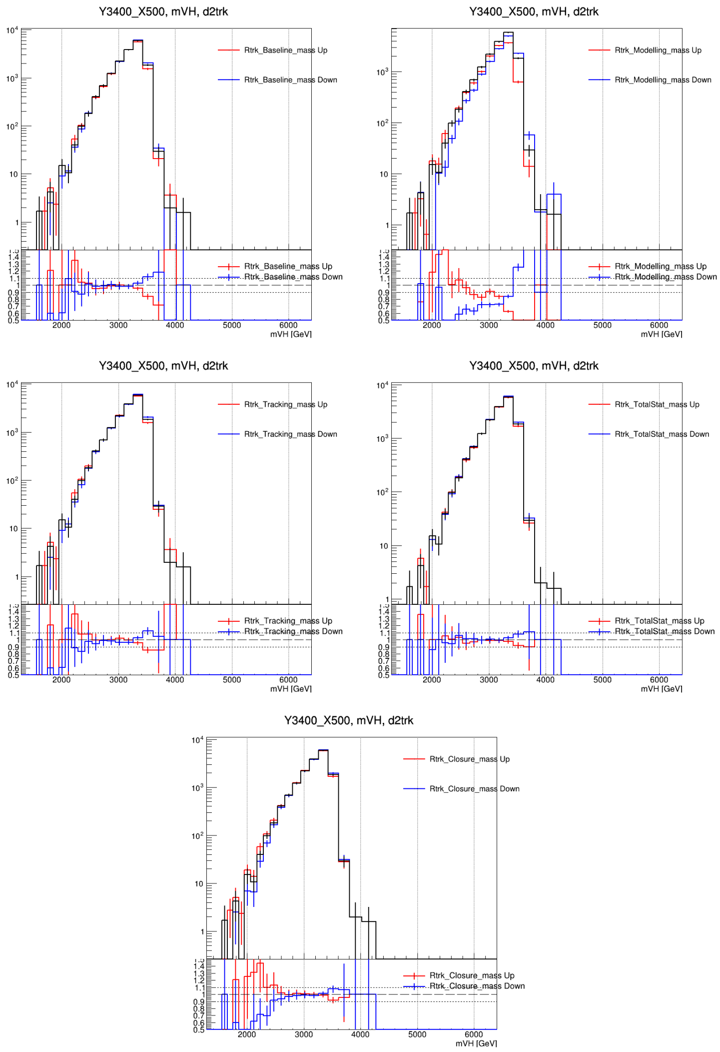


Figure 4.51: m_{JJ} distributions of the five RTrack jet CP variations (in order: Baseline, Modeling, Tracking, TotalStat, Closure) for an example $Y \to XH$ signal point of (m_Y=3400, m_X=500) GeV, on mass.

$H \to b\bar{b}$ uncertainties

Scale factors (SFs) must be applied to calibrate the $H \to b\bar{b}$ tagger and ensure the same performance from signal MC to the data-driven background estimation.
These are generated with dedicated analyses performed by the flavour tagging combined performance group. A proof of principle analysis that derives and validates these SFs for the $H \to b\bar{b}$ signal using $Z \to bb$ events as a proxy was published in 2021 [6]. There a previous training of the network was used, therefore scale factors (SFs) for this analysis have been computed to match the $H \to b\bar{b}$ tagger efficiency between data and simulation are included that scale the signal template, and their uncertainties are similarly propagated to the signal normalization. These SFs are computed using the methodology from Ref. [38], with the substitution of an updated $D_{H_{bb}}$ that includes the eta-reweighting of inputs. They are binned in large-R jet p_T, where the highest p_T bin SF is extrapolated to cover the upper end of the p_T regime probed by the analysis selection, and shown in Table 4.4.

p_T[GeV]	450-500	500-600	600-1000
SF	1.157	1.151	0.847
Total uncertainty	+0.268, -0.333	+0.317, -0.364	+0.218, -0.328

Table 4.4: Scale factors for the $H \to b\bar{b}$ tagger and their uncertainties, in bins of large-R jet p_T.

Uncertainties on the D_2^{trk} distribution and from the signal interpolation procedure are found to be negligible therefore they were not considered in the fit procedure.

4.6.3 Data-Driven Background Systematics

Variations on the background shape are derived in a custom way from three sources: the chosen of a specific training region, its finite size and the extrapolation procedure.
All background uncertainties are derived using an m_{JJ} shape that is inclusive in m_X, and applied to each m_X category fully correlated across m_{JJ} bins. This choice was made to avoid to be affected by the limited statistics in single $m_X$bins, supported by the evidence that studies comparing the uncertainty variations across bins of $m_X$did not indicate significant differences, as shown below in this Section.

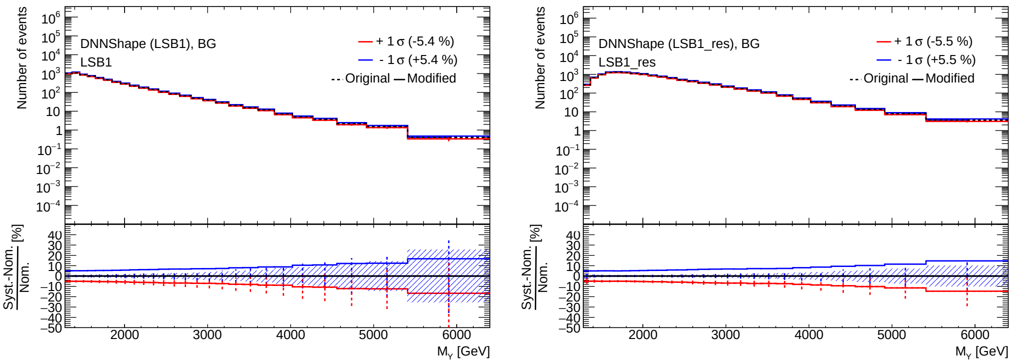
Training region uncertainty

The first uncertainty is associated to the dependence of the DNN model on the regions chosen for the training, since the predictions in the SR may be different if the region used for training changes. This uncertainty, here called *DNNShape* uncertainty, has been calculated by training the DNN in an alternate region of $165 < m_H < 200$ GeV, to account for potential variations in obtained weights due to differences in phase space.
The alternative training region has approximately the same statistics and tagging efficiency as the nominal training region, helping to isolate the effect of the particular DNN model on the obtained reweights.
Up and down variations are defined by symmetrizing the shape difference in m_{JJ} between the two different models, creating an effect of $\mathcal{O}(1-10)\%$ across the distribution.
Figure 4.52 shows the resulting variation of this study applied to the LSB validation region inclusively in m_X . For the final fit it is calculated in CR0.

Bootstrap procedure for the statistical uncertainty on the background shape

Another DNN variation is built to account for the finite statistics of the training sample and the random initialization of the network weights. These effects become asymptotically negligible if the training sample is large enough and as the number of training epochs increases. In this analysis both the conditions are quite satisfied but, in order to have an estimate of the size of this uncertainty a bootstrap procedure [100] has been adopted. This uncertainty corresponds to a $\mathcal{O}(1)\%$ effect across m_Y.

A set of 100 bootstrap networks are trained, each time varying the training dataset by re-sampling it with replacement. In practice sampling with replacement is obtained by simply weighting each event in the dataset with a Poisson variable with



(a) (b)

Figure 4.52: NN-based background re-weighting shape uncertainty in the LSB inclusive in m_X for both the Two-prong merged (left) and resolved (right) selection criteria.

mean 1[3]. In this way the b-th network (b is the index running over the N bootstraps) is trained on a training dataset each time a bit different from the previous one. Repeating the training N times allows to simultaneously take into account both the weights initialization and the dataset effects.
Two additional templates are then defined with the median weight for each event, plus or minus half of the interquartile (IQR) range, defining the upper and lower symmetric error bands.

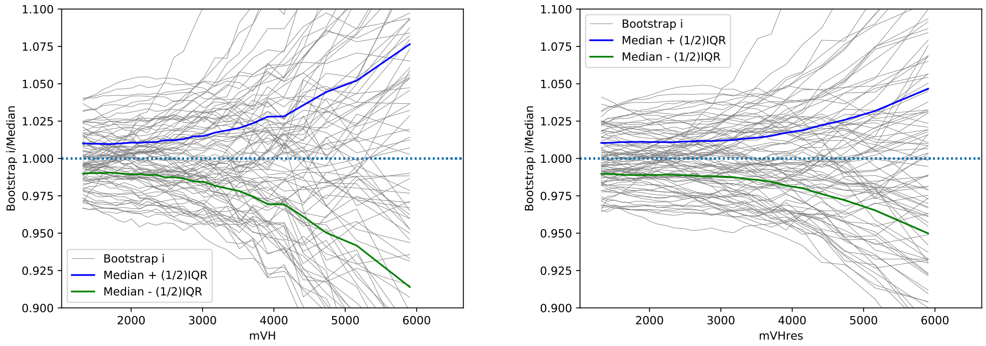
The N different training does not only affect the shape, but also the normalization of the template since for the b-th bootstrap it is defined as $M^b = \frac{k}{\Sigma_i w_i^b}$, where k is a general normalization, and depends on the particular set of weights.
The normalization and the shape effects are treated as separate, calculating the median value and the quartiles for the N normalization factors as well.Therefore the nominal template is normalized with the median value of the normalizations and the up and down variations with the median normalization +/- half of their IQR[4].

In figure 4.53 an estimate of the magnitude of the error calculated with N=100

[3] On each re-sampling with replacement of the dataset the i-th event can have multiplicity n_i, described by a Binomial distribution with the number of total extractions equal to N (the dataset size) and the probability to be extracted equal to $1/N$. If N becomes large enough, the Binomial distribution approximates the Poissonian distribution with mean equal to $N \cdot 1/N = 1$

[4] In more details, before applying the normalization factor to each template, the latter is re-scaled to the yields of the nominal histogram

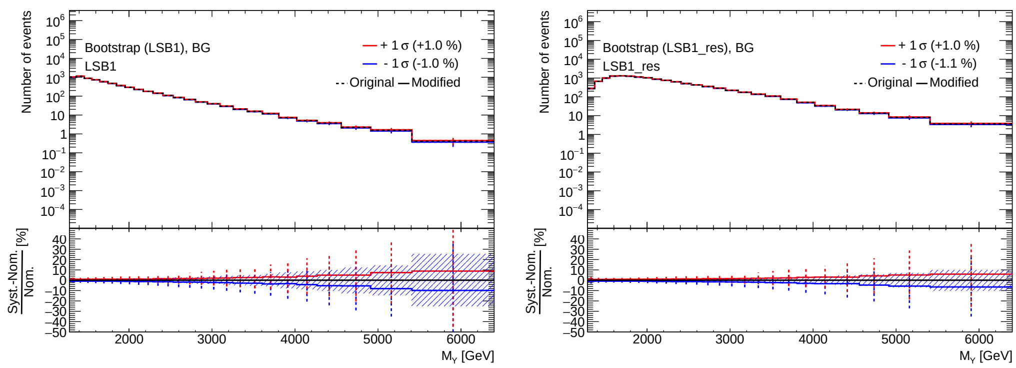
bootstrap networks is shown, for m_{JJ}, in merged and resolved CR0.
Figure 4.54 shows the resulting bootstrap uncertainty applied to the LSB validation



(a) (b)

Figure 4.53: Ratios among each of the N bootstraps histograms and the nominal one, for the m_{XH} shape calculated for the Two-prong merged (a) and resolved (b) X reconstruction. Blue and green ratios are for up and down histograms.

region inclusively in m_X for both the merged and resolved selection criteria.

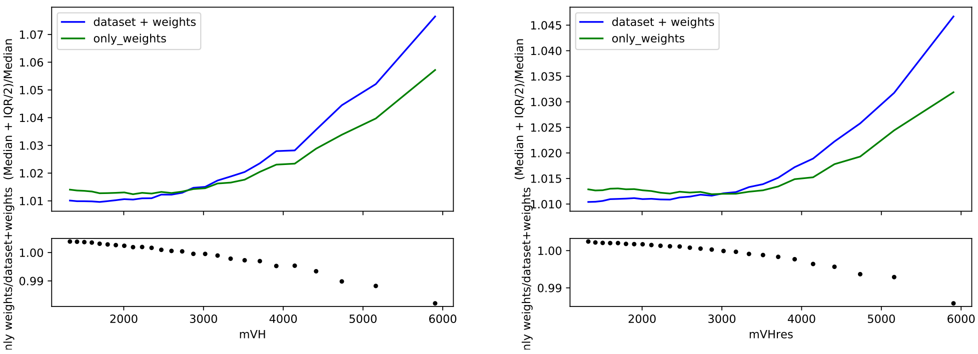


(a) (b)

Figure 4.54: NN-based bootstrap uncertainty in the LSB inclusive in m_X for both the merged (left) and resolved (right) selection criteria.

In order to estimate the size of the finite statistics effect with respect to that of the casual weight initialization, the bootstrap procedure has been repeated, this time fixing the training dataset and allowing only the weights initialization to vary. Results are shown in Figure 4.55 where the upper error band obtained with the

only weights variations case is compared with the *dataset+weights*. Comparisons are performed in CR0 (m_H in [75, 145] GeV), inclusively in m_X since there are no reasons why the error contributions should depend on m_X.
From this study it can be seen that the dataset impact is very small, because almost all the training variations can be addressed to the weights initialization and the stochastic loss optimization.



(a) (b)

Figure 4.55: Up variations compared between the *only weights variations* case and the *dataset+weights* error band, in CR0, both in Two-prong merged and resolved categories, inclusively in m_X. In the bottom plots, the ratio between the green and the blue curves is shown.

Non-closure uncertainty

Lastly, a non-closure uncertainty is included to cover modeling discrepancies that may arise from extrapolating weights derived from the NN training in the HSB to the LSB, and subsequently to the SR.
It is defined by the symmetrized shape difference between the data and predicted background in the LSB. In order to not be sensitive to statistical fluctuations in the LSB, a smoothing is applied to the variation where it is re-binned to reduce the relative statistical uncertainty.
The non-closure is negligible for low m_{JJ} and rises to $\mathcal{O}(10)\%$ in the m_{JJ} tails.
Figure 4.56 shows these symmetrized deviations inclusively in m_X . The ratio plot includes the unsmoothed variation (dotted) as well as the smoothed variation (solid).

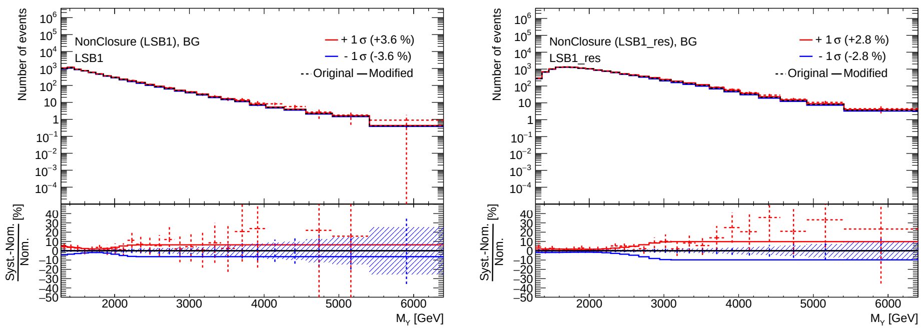


Figure 4.56: Non-closure uncertainty on m_{JJ} computed inclusively in m_X for both the merged (left) and resolved (right) selection criteria.

4.6.4 m_X Exclusive vs. Inclusive Shapes

In this Section comparisons between background systematics calculated in each $m_X$window are shown, for the three uncertainties discussed above. No significant differences are shown, motivating the adoption of an m_X-inclusive uncertainty.

Figures 4.57 to 4.59 show the ratio between the exclusively vs. inclusively computed DNNShape uncertainty with the Two-prong merged requirement applied, in some of the m_X window produced.

Figures 4.60 to 4.62 show the ratio between the exclusively vs. inclusively computed DNNShape uncertainty with the Two-prong resolved selection requirement applied, in some of the m_X window produced.

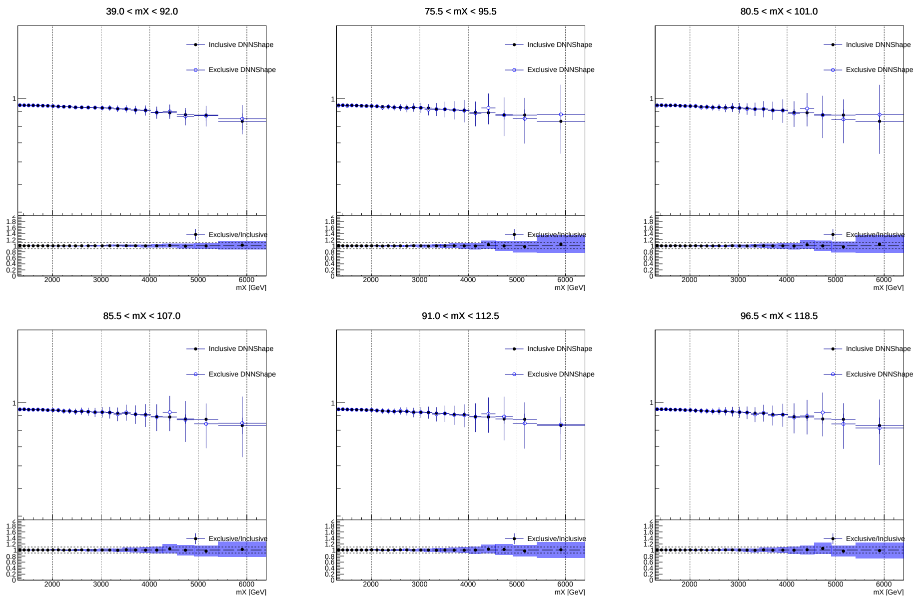


Figure 4.57: Ratio of the exclusively vs. inclusively determined DNNShape uncertainty in Two-prong merged region for m_X windows between 39.0 and 118.5 GeV.

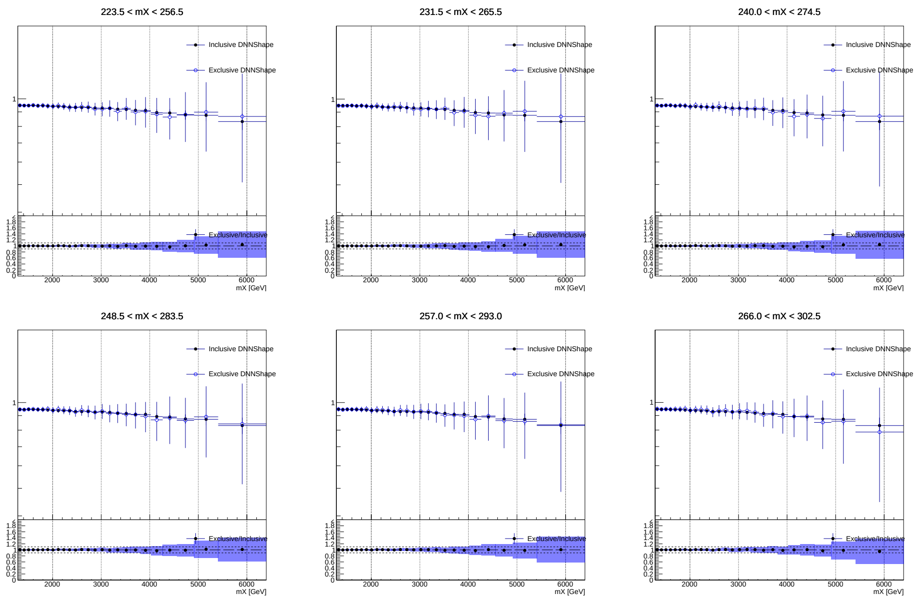


Figure 4.58: Ratio of the exclusively vs. inclusively determined DNNShape uncertainty in Two-prong merged region for m_X windows between 223.5 and 302.5 GeV.

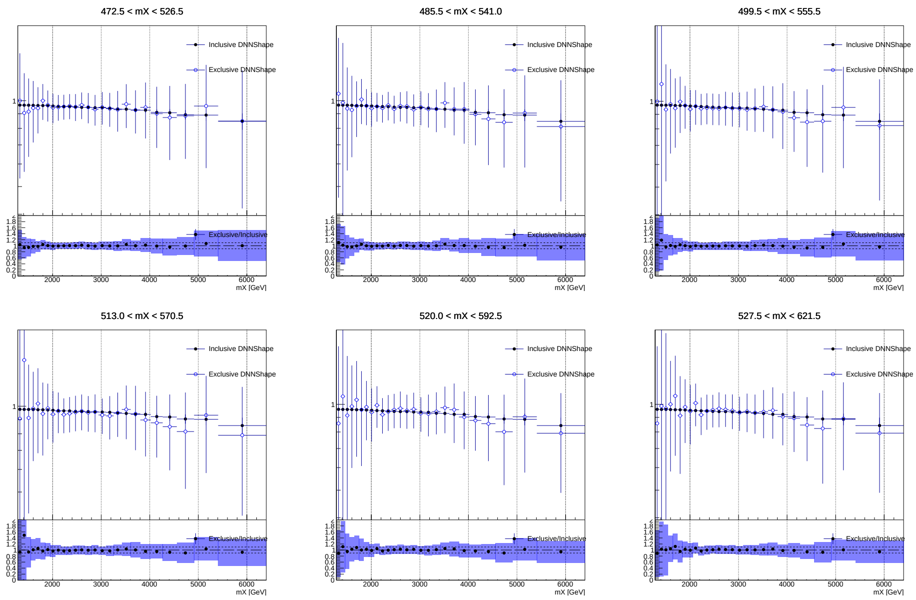


Figure 4.59: Ratio of the exclusively vs. inclusively determined DNNShape uncertainty in Two-prong merged region for m_X windows between 472.5 and 621.5 GeV.

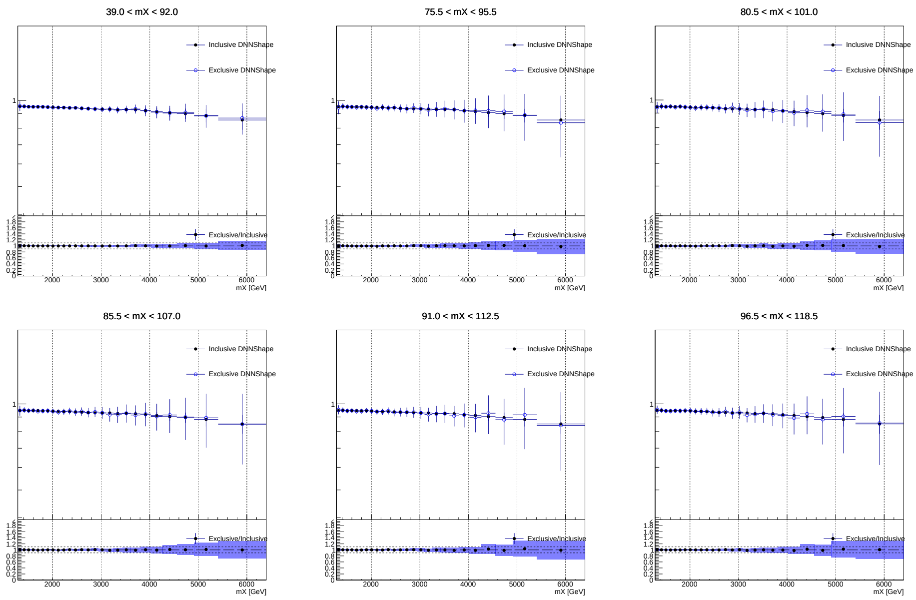


Figure 4.60: Ratio of the exclusively vs. inclusively determined DNNShape uncertainty in Two-prong resolved region for m_X windows between 39.0 and 118.5 GeV.

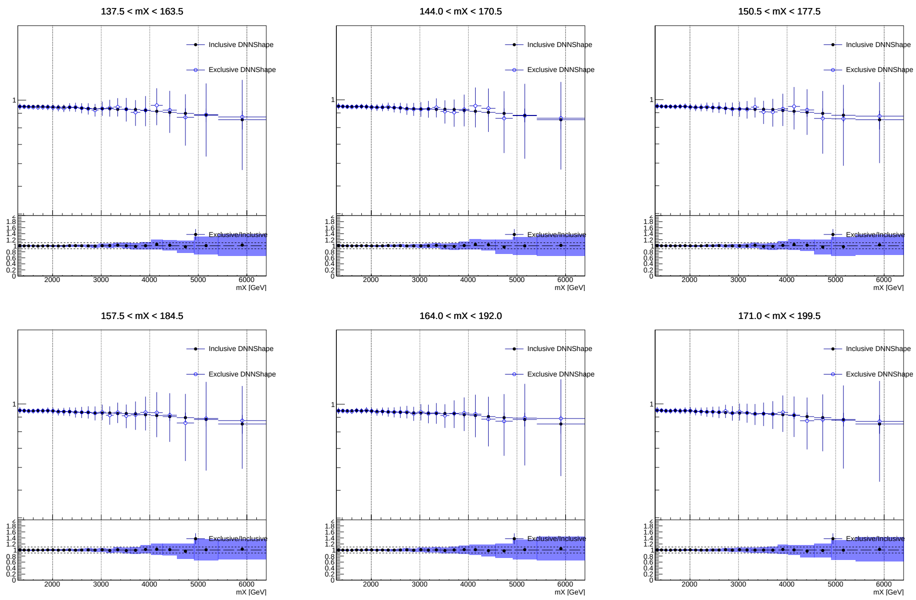


Figure 4.61: Ratio of the exclusively vs. inclusively determined DNNShape uncertainty in Two-prong resolved region for m_X windows between 137.5 and 199.5 GeV.

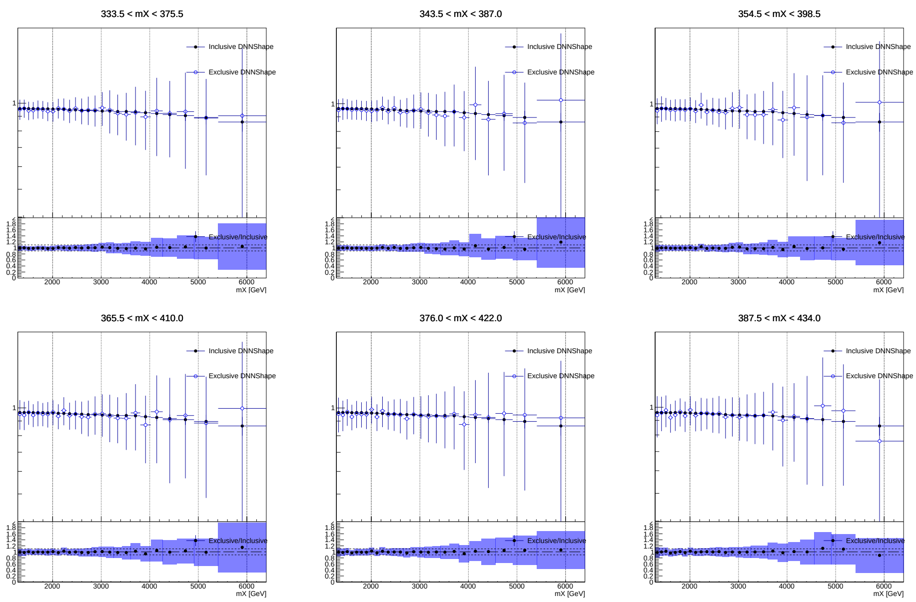


Figure 4.62: Ratio of the exclusively vs. inclusively determined DNNShape uncertainty in Two-prong resolved region for m_X windows between 333.5 and 434.0 GeV.

Figures 4.63 to 4.65 show the ratio between the exclusively vs. inclusively computed DNN Bootstrap's up variation with the Two-prong merged requirement applied, in some of the m_X window produced.

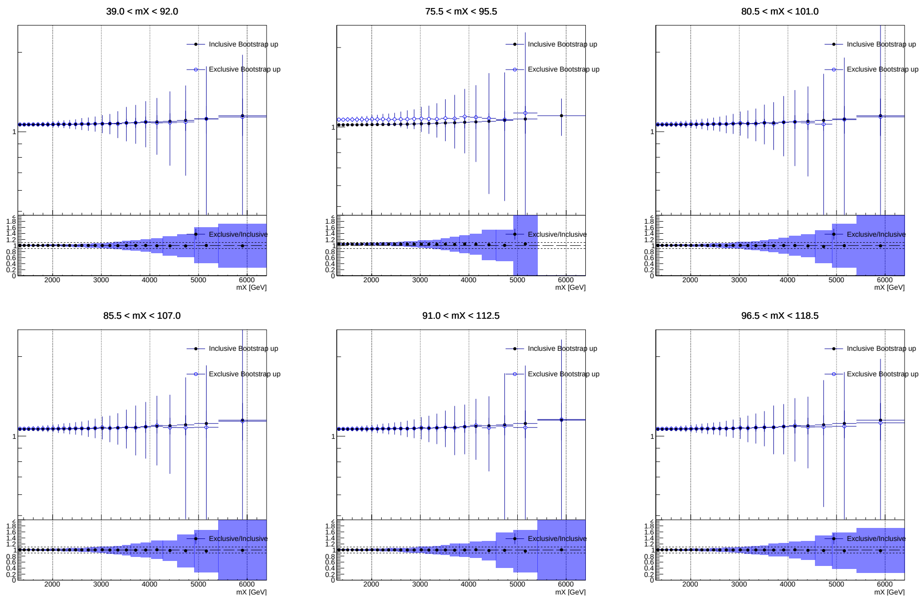


Figure 4.63: Ratio of the exclusively vs. inclusively determined DNN Bootstrap's up variation in Two-prong merged region for m_X windows between 39.0 and 118.5 GeV.

Figures 4.66 to 4.68 show the ratio between the exclusively vs. inclusively computed DNN Bootstrap's up variation with the Two-prong resolved selection requirement applied, in some of the m_X window produced.

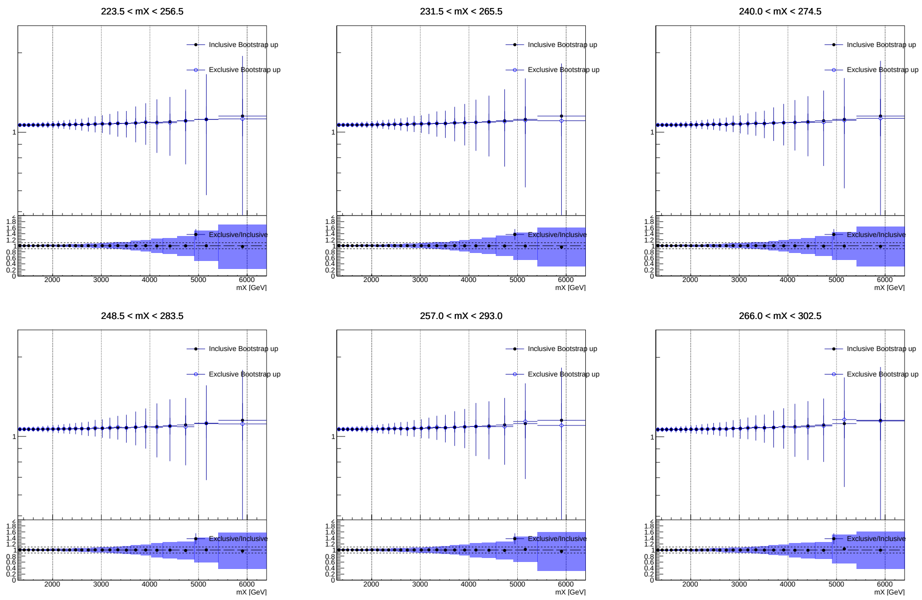


Figure 4.64: Ratio of the exclusively vs. inclusively determined DNN Bootstrap's up variation in Two-prong merged region for m_X windows between 223.5 and 302.5 GeV.

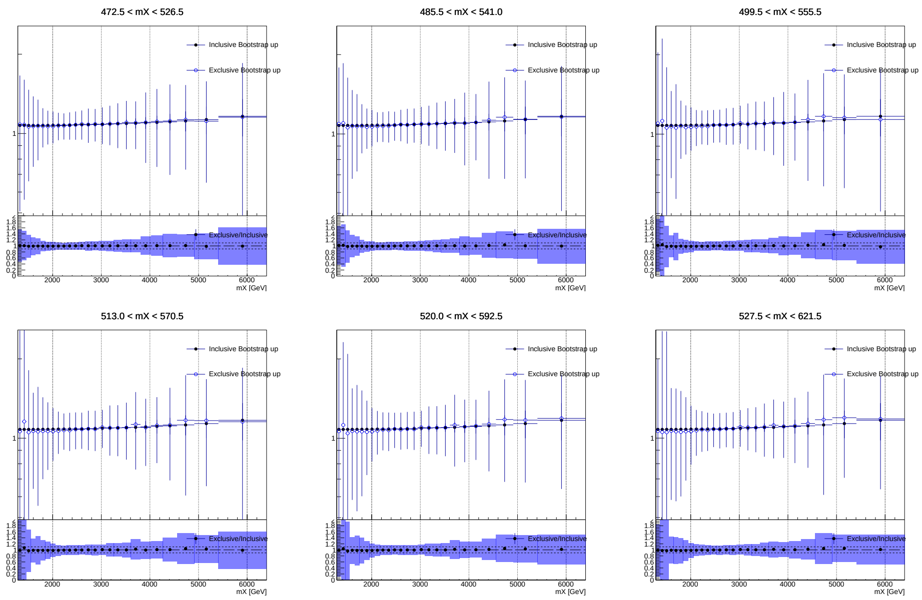


Figure 4.65: Ratio of the exclusively vs. inclusively determined DNN Bootstrap's up variation in Two-prong merged region for m_X windows between 472.5 and 621.5 GeV.

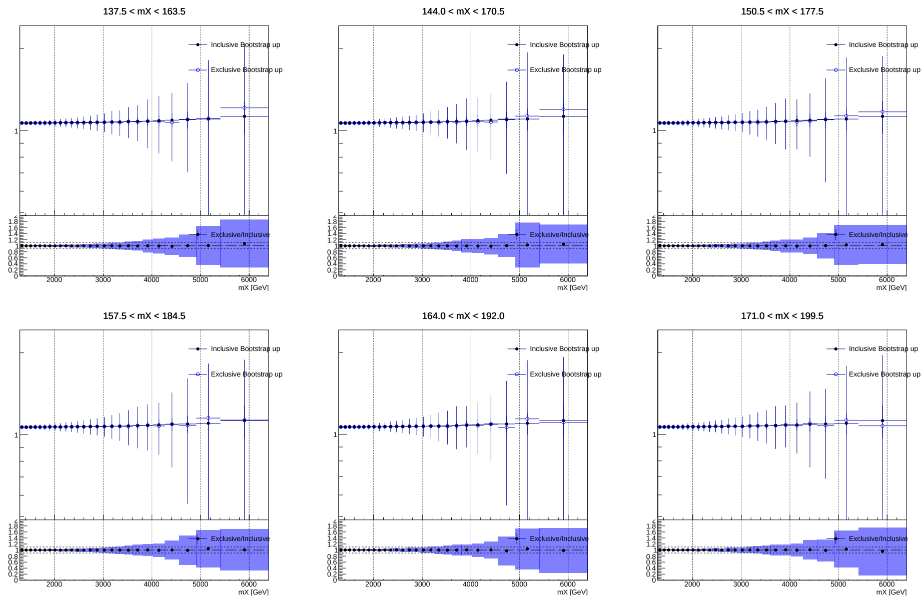


Figure 4.66: Ratio of the exclusively vs. inclusively determined DNN Bootstrap's up variation in Two-prong resolved regionrequirement applied for m_X windows between 137.5 and 199.5 GeV.

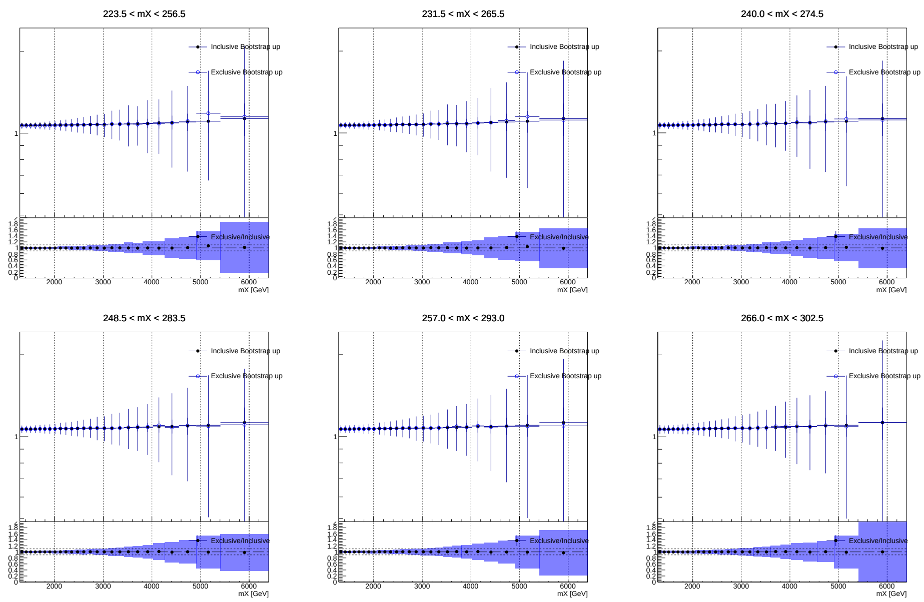


Figure 4.67: Ratio of the exclusively vs. inclusively determined DNN Bootstrap's up variation in Two-prong resolved regionrequirement applied for m_X windows between 223.5 and 302.5 GeV.

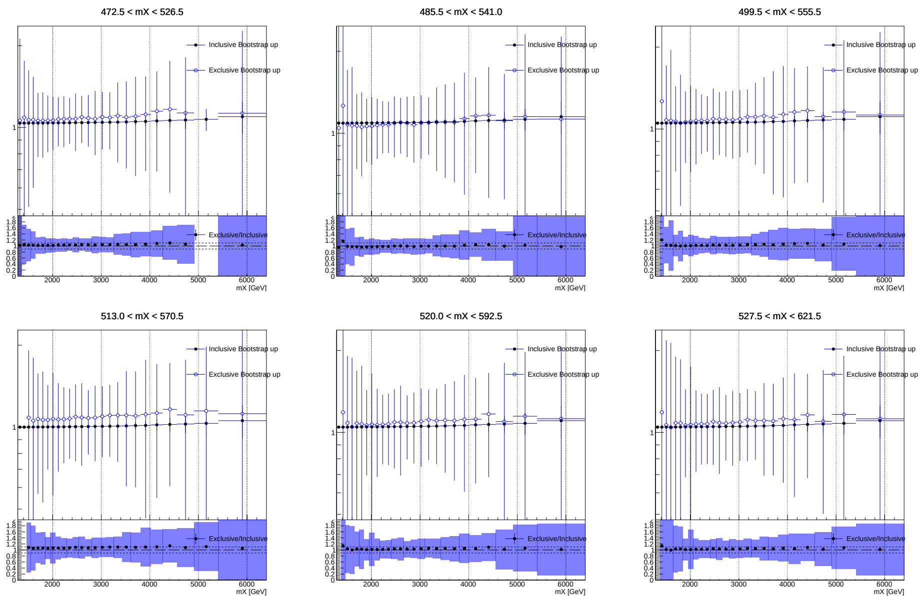


Figure 4.68: Ratio of the exclusively vs. inclusively determined DNN Bootstrap's up variation in Two-prong resolved regionrequirement applied for m_X windows between 472.5 and 621.5 GeV.

Figures 4.69 to 4.71 show the ratio between LSB1 and reweighted LSB0 data in the LSB validation region, as the Non-closure uncertainty is calculated, with the Two-prong merged requirement applied, in some of the m_X window produced.

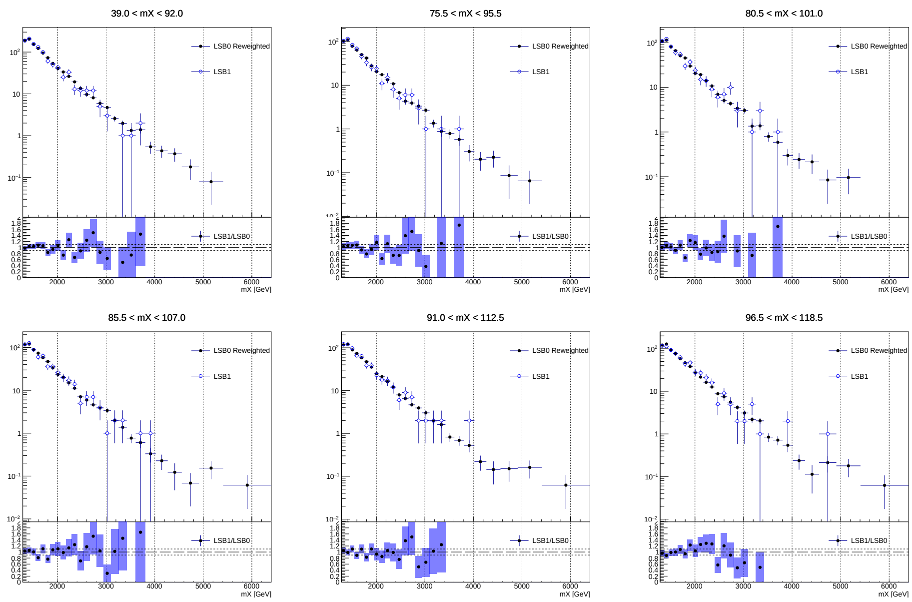


Figure 4.69: LSB1/LSB0 ratio in Two-prong merged region for m_X windows between 39.0 and 118.5 GeV.

Figures 4.72 to 4.74 show the ratio between LSB1 and reweighted LSB0 data in the LSB validation region, as the Non-closure uncertainty is calculated, with the Two-prong resolved requirement applied, in some of the m_X window produced.

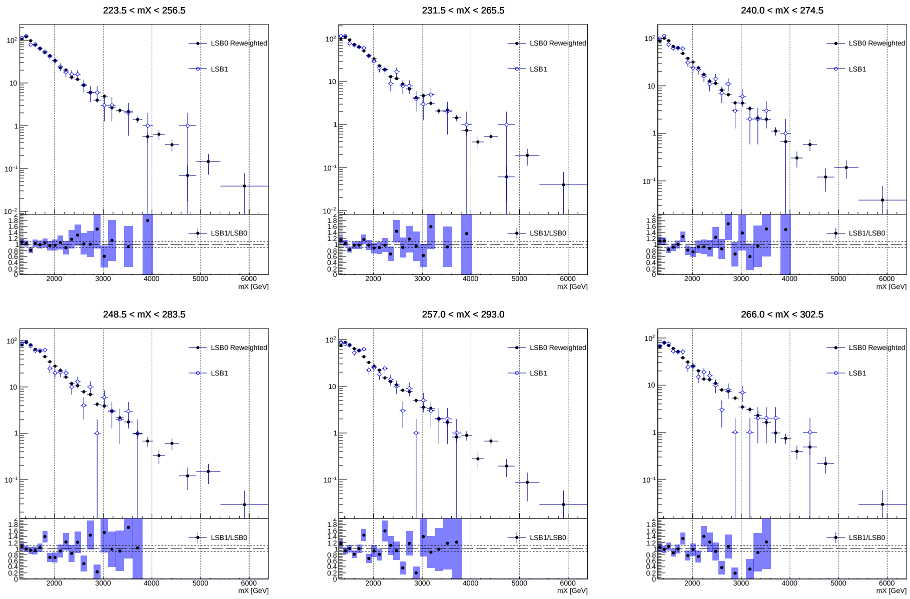


Figure 4.70: LSB1/LSB0 ratio in Two-prong merged region for m_X windows between 223.5 and 302.5 GeV.

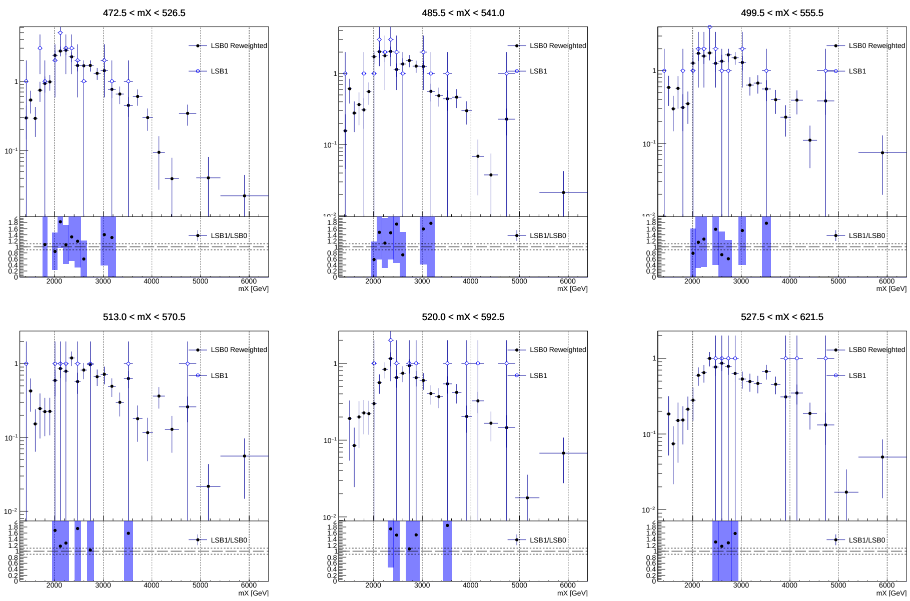


Figure 4.71: LSB1/LSB0 ratio in Two-prong merged region for m_X windows between 472.5 and 621.5 GeV.

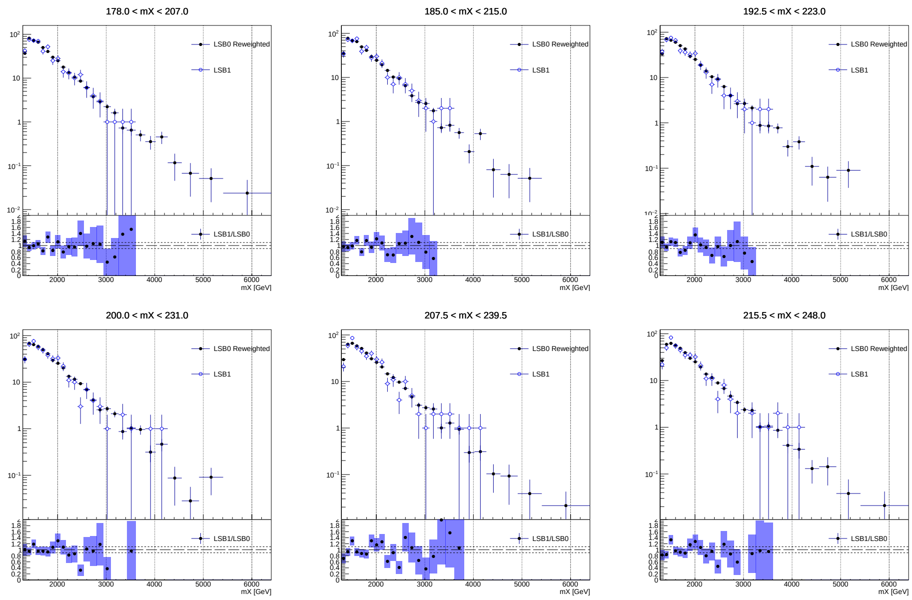


Figure 4.72: LSB1/LSB0 ratio with the resolved selection requirement applied for m_X windows between 178.0 and 248.0 GeV.

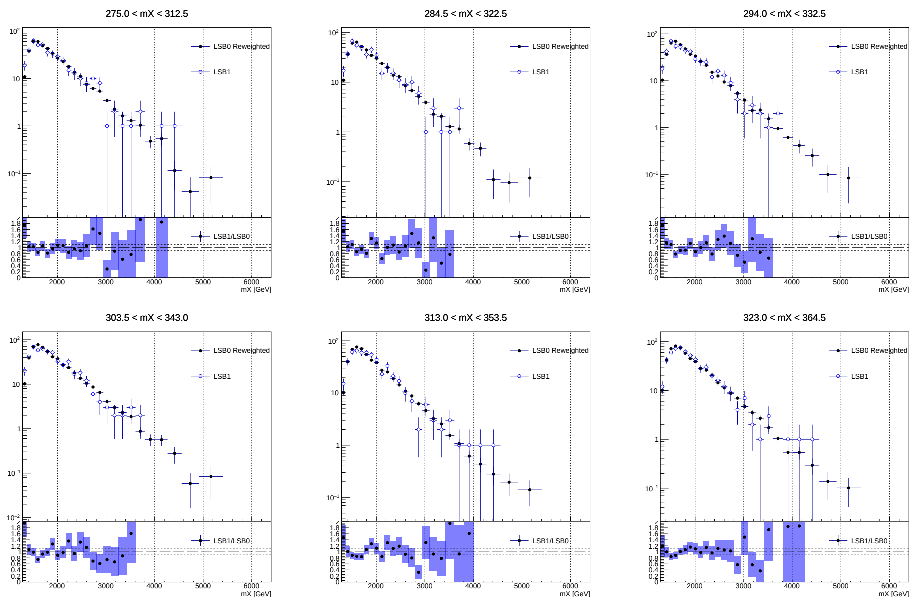


Figure 4.73: LSB1/LSB0 ratio with the resolved selection requirement applied for m_X windows between 275.0 and 364.5 GeV.

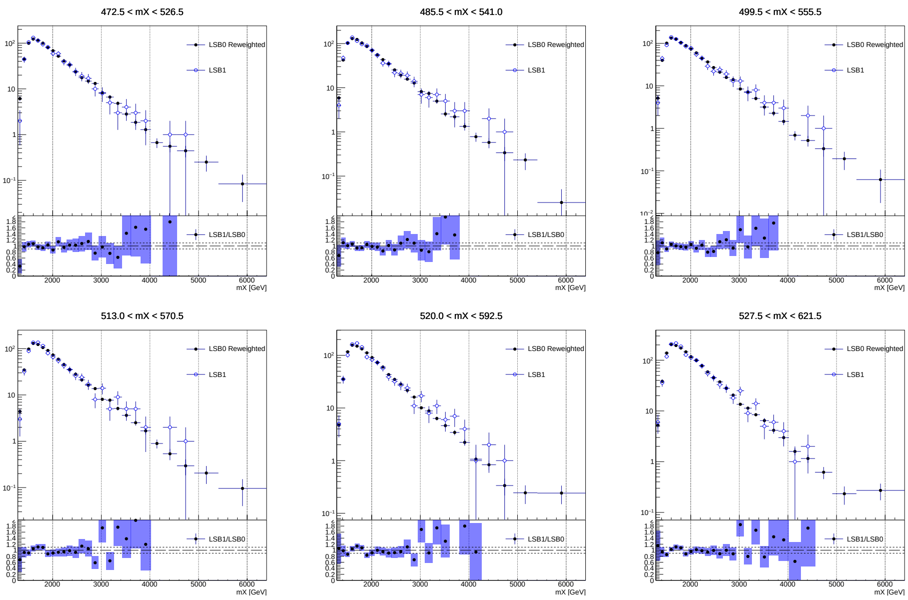


Figure 4.74: LSB1/LSB0 ratio with the resolved selection requirement applied for m_X windows between 472.5 and 621.5 GeV.

Figures 4.75 to 4.77 show the ratio between LSB1 and reweighted LSB0 data in the LSB validation region, as the Non-closure uncertainty is calculated, with the Anomaly Score on the X candidate > 0.5 requirement applied, in some of the m_X window produced.

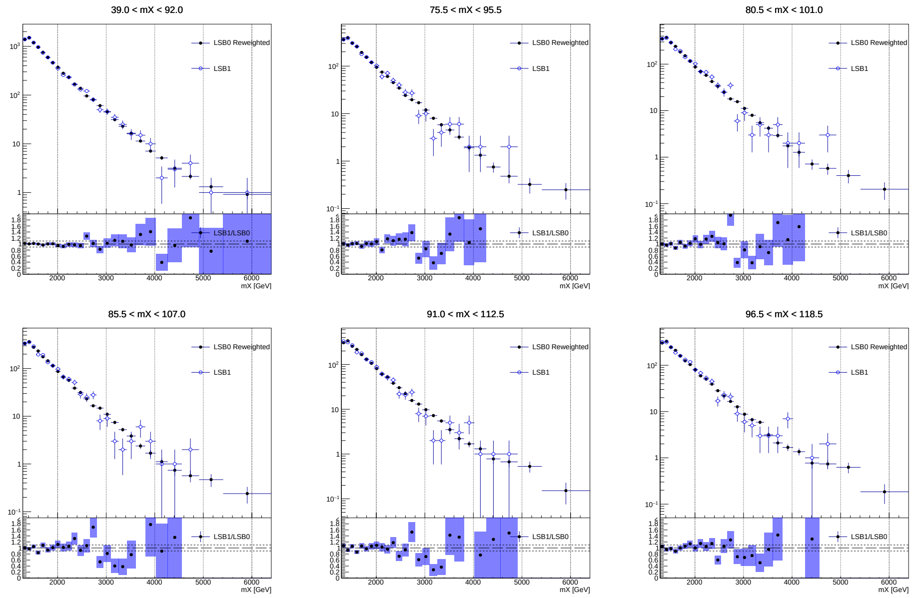


Figure 4.75: LSB1/LSB0 ratio in Anomaly region for m_X windows between 39.0 and 118.5 GeV.

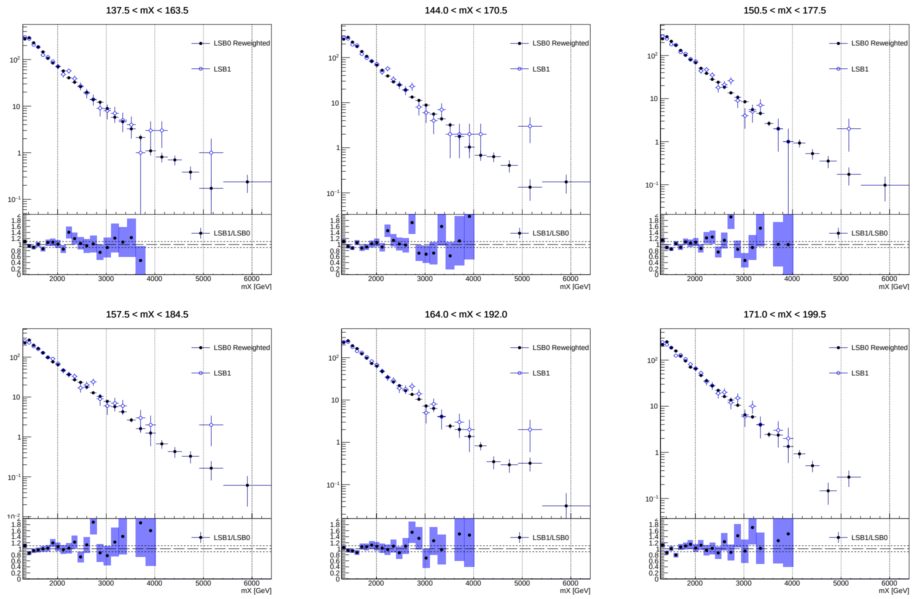


Figure 4.76: LSB1/LSB0 ratio in Anomaly region for m_X windows between 137.5 and 199.5 GeV.

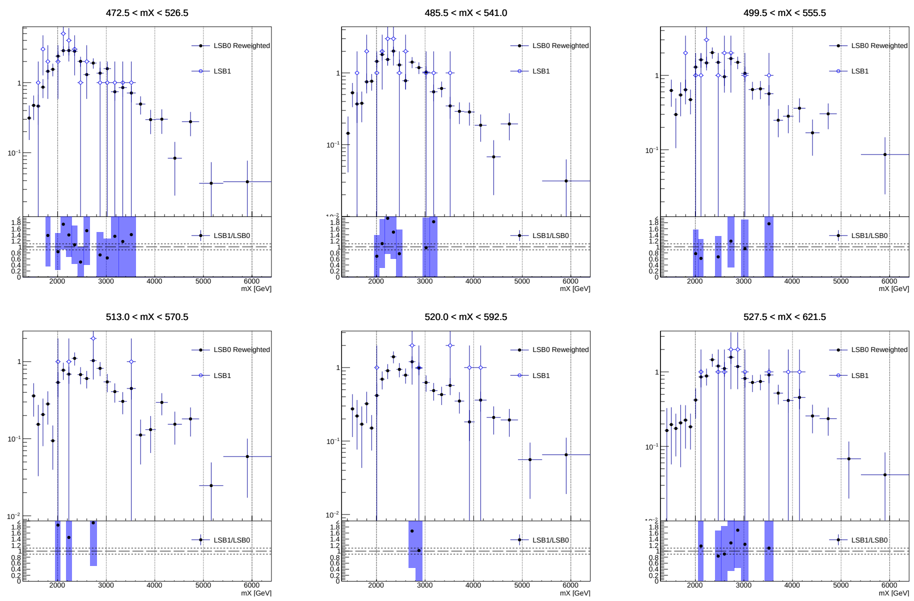


Figure 4.77: LSB1/LSB0 ratio in Anomaly region for m_X windows between 472.5 and 621.5 GeV.

4.7 Statistical analysis

The statistical analysis consisted in performing hypothesis testing in the SRs, for the compatibility of the data with both the background-only and the signal-plus-background hypotheses. The observable that is fit is the m_{JJ} distribution of the data in the SR. This fit is repeated several times in overlapping bins of the *X* candidate mass.
Two different approaches were followed: a search for excesses on data with respect to the background-only hyposthesis was performed in the Anomaly region, while both the background-only and the signal-plus-background hypotheses were tested in the Two-prong merged and resolved region for each $Y \rightarrow XH$ signal model and 95% confidence level upper limits on the signal cross section were set using a modified frequentist method (CL_s) [104].

4.7.1 Binning optimization

Both fit strategies implement a binning of m_Y and m_X that is informed by the linear fits of expected signal mass resolution. The binning is chosen based on the m_Y and m_X resolution, with modifications based on available statistics.
In the m_Y spectrum, bins are widened at higher values of m_Y to ensure that at least one event is available in each bin when applied to LSB1, where the background Non-closure uncertainty is derived (see Section4.6.3). This is performed by adding an exponential curve to the linear fit shown in Figure 4.45 such that the resulting binning satisfies the statistical requirement while remaining monotonically increasing. An upper limit of 6.4 TeV is chosen such that the choice of binning provides sensitivity to Y masses up to 6 TeV.

In the m_X spectrum, mass windows must be chosen such that enough statistics are available in order to have a robust fit and non-zero bins in the non-closure uncertainty. An initial set of mass windows is generated from the linear fit shown in Figure 4.45. The first window is chosen to have a bin center of m_X = 65 GeV, the lowest generated signal X mass, and its width chosen to be twice the resolution obtained from the fit. Subsequent windows are generated in the same way, with their bin centers chosen to be higher than the previous window's bin center by a value equal to half of that window's resolution. The edges of high-mass windows are expanded symmetrically based on statistics in LSB1, until at least 10 events

are present in each final m_X bin. In order to protect against duplicate windows, all resulting windows which share at least one bin edge are replaced by a single window encompassing their union.
Finally, to accommodate the highest generated signal mass point at m_X =3000 GeV, the final bin's high edge is increased to a value of 3200 GeV.

The final binning choice leads is the following, common to both the Anomaly and Two-prong regions.

m_Y bins: [1282, 1368, 1460, 1551, 1648, 1747, 1849, 1952, 2061, 2172, 2289, 2409, 2533, 2664, 2801, 2945, 3097, 3256, 3424, 3608, 3805, 4021, 4267, 4557, 4912, 5409, 6400]

m_X bins: [[39.0, 92.0], [75.5, 95.5], [80.5, 101.0], [85.5, 107.0], [91.0, 112.5], [96.5, 118.5], [102.0, 124.5], [107.5, 130.5], [113.0, 137.0], [119.0, 143.5], [125.5, 150.0], [131.5, 156.5], [137.5, 163.5], [144.0, 170.5], [150.5, 177.5], [157.5, 184.5], [164.0, 192.0], [171.0, 199.5], [178.0, 207.0], [185.0, 215.0], [192.5, 223.0], [200.0, 231.0], [207.5, 239.5], [215.5, 248.0], [223.5, 256.5], [231.5, 265.5], [240.0, 274.5], [248.5, 283.5], [257.0, 293.0], [266.0, 302.5], [275.0, 312.5], [284.5, 322.5], [294.0, 332.5], [303.5, 343.0], [313.0, 353.5], [323.0, 364.5], [333.5, 375.5], [343.5, 387.0], [354.5, 398.5], [365.5, 410.0], [376.0, 422.0], [387.5, 434.0], [399.0, 446.5], [410.5, 459.0], [422.5, 472.0], [434.5, 485.5], [447.0, 499.0], [460.0, 512.5], [472.5, 526.5], [485.5, 541.0], [499.5, 555.5], [513.0, 570.5], [520.0, 592.5], [527.5, 621.5], [537.5, 659.0], [546.5, 711.0], [581.0, 3200.0]]

4.7.2 Anomaly Region

As the anomaly score is not strongly dependent on any particular signal model, and the Two-prong regions have been designed to give better cross section limits to the specified 2-prong signal grid, no signal interpretation has been done in the anomaly region. Results have been provided in the form of background-only p-values, assessing the compatibility of the signal region data with a background-only hypothesis, using the BumpHunter [47] algorithm.
To validate the fit procedure, spurious signal and signal injection tests are performed in blinded data regions.
A threshold value of 0.01 for the p-values is used as a metric for a significant excess with respect to background.

Spurious signal tests

To ensure that the cut on the anomaly score does not produce spurious excesses where no $Y \to XH$ signal is present, the BUMPHUNTER (BH) fit procedure has been applied for background-only samples in the LSB.
The DNN weights determined for the background reweighting have been applied to the shape distribution in LSB0 after the AS cut to provide a background estimation for LSB1 data. The LSB0 distributions have been further rescaled so that they matched the yield in LSB1.
The BUMPHUNTER algorithm took in input these distributions, in each m_X bin, and provided a p-value measuring the compatibility of LSB1 with LSB0 reweighed data.
Table 4.5 summarizes these p-values, where no significant excesses are found deviating from the background only model. Figure 4.78 displays this table as a histogram, to demonstrate that there is no significant bias of the expected compatibility with respect to m_X.

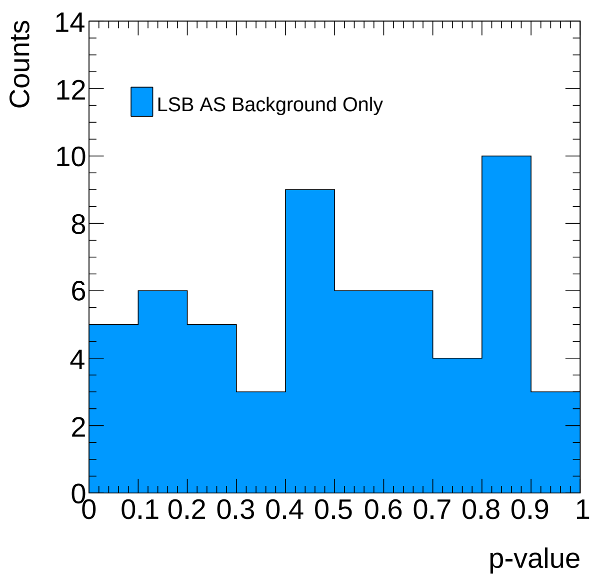


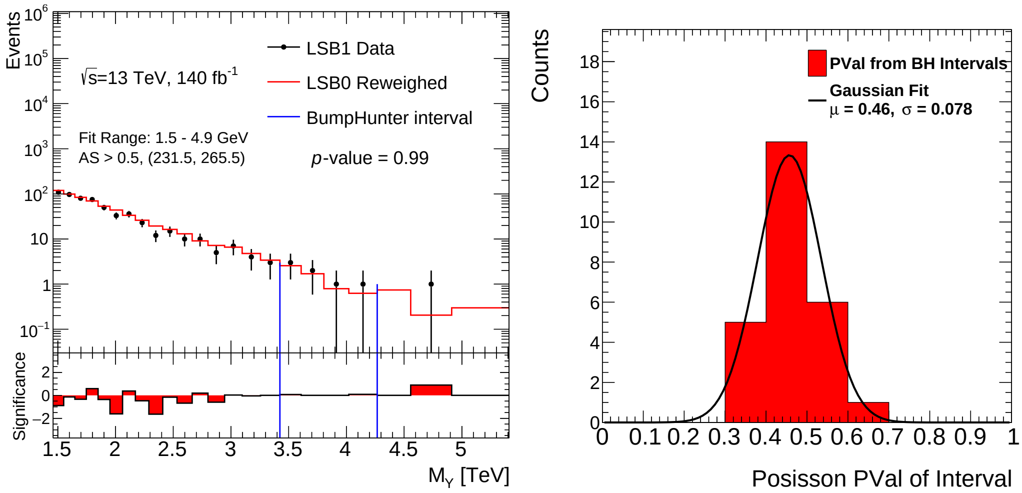
Figure 4.78: Histogram representation of Table 4.5, i.e. the distribution of p-values obtained for each m_X bin background only fit. As these approximate a flat distribution between 0 and 1, no bias in background-to-data compatibility is observed.

Figure 4.79 shows plots for the sample m_X bin (231.5, 265.5) GeV presented in the table, along with the distribution of p-values obtained for all the BH intervals tested

Table 4.5: BumpHunter p-values in each m_X bin from spurious signal tests. No significant excess (p-value < 0.01) is found.

X Bin	p-value	BH Interval	X Bin	p-value	BH Interval
(39.0, 92.0)	0.434	(2533.0, 2801.0)	⋮	⋮	⋮
(75.5, 95.5)	0.248	(2172.0, 2801.0)	(266.0, 302.5)	0.586	(2172.0, 2533.0)
(80.5, 101.0)	0.055	(2172.0, 2801.0)	(275.0, 312.5)	0.831	(1460.0, 1849.0)
(85.5, 107.0)	0.302	(2289.0, 2801.0)	(284.5, 322.5)	0.486	(2289.0, 2533.0)
(91.0, 112.5)	0.553	(3424.0, 4021.0)	(294.0, 332.5)	0.463	(2289.0, 2533.0)
(96.5, 118.5)	0.138	(3608.0, 4021.0)	(303.5, 343.0)	0.615	(2289.0, 3424.0)
(102.0, 124.5)	0.591	(3424.0, 4021.0)	(313.0, 353.5)	0.413	(2801.0, 4557.0)
(107.5, 130.5)	0.143	(3424.0, 4021.0)	(323.0, 364.5)	0.723	(2289.0, 2533.0)
(113.0, 137.0)	0.222	(2172.0, 2409.0)	(333.5, 375.5)	0.575	(3256.0, 3805.0)
(119.0, 143.5)	0.284	(2172.0, 2533.0)	(343.5, 387.0)	0.501	(2061.0, 2409.0)
(125.5, 150.0)	0.043	(3424.0, 5409.0)	(354.5, 398.5)	0.712	(1849.0, 2172.0)
(131.5, 156.5)	0.03	(2172.0, 2664.0)	(365.5, 410.0)	0.472	(1551.0, 1747.0)
(137.5, 163.5)	0.137	(2172.0, 2409.0)	(376.0, 422.0)	0.463	(1952.0, 2172.0)
(144.0, 170.5)	0.076	(2172.0, 2801.0)	(387.5, 434.0)	0.857	(2664.0, 3097.0)
(150.5, 177.5)	0.15	(2533.0, 2801.0)	(399.0, 446.5)	0.855	(2061.0, 3097.0)
(157.5, 184.5)	0.213	(2533.0, 2801.0)	(410.5, 459.0)	0.403	(3256.0, 3608.0)
(164.0, 192.0)	0.317	(2664.0, 2945.0)	(422.5, 472.0)	0.641	(1460.0, 1648.0)
(171.0, 199.5)	0.55	(1648.0, 2172.0)	(434.5, 485.5)	0.636	(1460.0, 1747.0)
(178.0, 207.0)	0.733	(3608.0, 4021.0)	(447.0, 499.0)	0.056	(1460.0, 1747.0)
(185.0, 215.0)	0.877	(3608.0, 4021.0)	(460.0, 512.5)	0.147	(1460.0, 1747.0)
(192.5, 223.0)	0.877	(3608.0, 4021.0)	(472.5, 526.5)	0.444	(1551.0, 1747.0)
(200.0, 231.0)	0.806	(1551.0, 1747.0)	(485.5, 541.0)	0.633	(2172.0, 2664.0)
(207.5, 239.5)	0.876	(2801.0, 3256.0)	(499.5, 555.5)	0.243	(2172.0, 2409.0)
(215.5, 248.0)	0.853	(1747.0, 1952.0)	(513.0, 570.5)	0.103	(2172.0, 2409.0)
(223.5, 256.5)	0.971	(1747.0, 1952.0)	(520.0, 592.5)	0.408	(3424.0, 4267.0)
(231.5, 265.5)	0.993	(3424.0, 4267.0)	(527.5, 621.5)	0.68	(2289.0, 3097.0)
(240.0, 274.5)	0.674	(3424.0, 4267.0)	(537.5, 659.0)	0.831	(2289.0, 2945.0)
(248.5, 283.5)	0.952	(3424.0, 4267.0)	(546.5, 711.0)	0.747	(2289.0, 2664.0)
(257.0, 293.0)	0.898	(1648.0, 1849.0)	(581.0, 3200.0)	0.35	(2409.0, 2664.0)
⋮	⋮	⋮			

in the fit. The normal distribution of interval p-values indicates that no systematic trend of spurious signal is present across the m_Y distribution. The largest observed deviation is in m_X bin (131.5, 156.5) with a p-value of 0.03, and is illustrated in figure 4.80a.



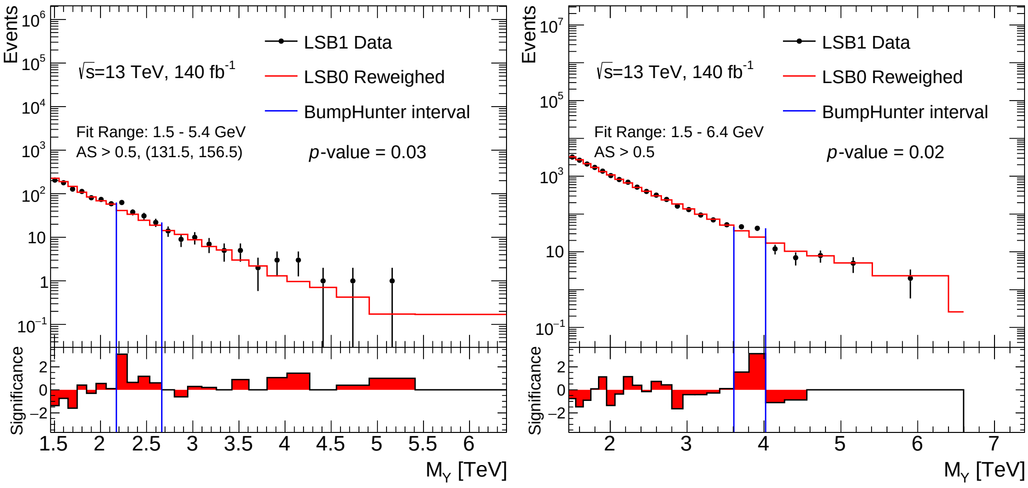
(a) (b)

Figure 4.79: (a) BUMPHUNTER spurious signal test plots in the sample m_X bin (231.5, 265.5) GeV. No significant excesses over the background are found. (b) corresponding histograms of p-value distributions for all intervals checked in the BH fit, fit to a Gaussian.

To account for signal region statistics, these studies were also performed with Asimov toys in the Anomaly SR. The CR0 is reweighed with DNN weights and toys generated bin-by-bin by sampling a Poisson distribution with mean equal to the value of the corresponding bin in the reweighed CR0 distribution. Table 4.6 summarizes these p-values and two sample BUMPHUNTER fits are shown in Figure 4.81.

Signal injection tests

Signal injection tests have also been performed to validate that BUMPHUNTER can pick out a signal injected over background in the LSB. The injected signals are scaled according to $S/\sqrt{B}$ significance to be at the 2σ and 0.5σ level.
The results of the signal injection are summarized in Figure 4.82. Figure 4.83 shows whether the m_Y value of the injected signal is within the determined BUMPHUNTER interval. This accounts for p-values corresponding to background fluctuations, and shows good sensitivity in the AS-selected data across the entire signal grid at the 2.0σ level, and to the highly boosted points at the 0.5σ level.
Sample BUMPHUNTER plots are shown in Figure 4.7.2 for both 2σ and 0.5σ signal



(a) (b)

Figure 4.80: BUMPHUNTER spurious signal test (a) in m_X bin (131.5, 156.5) and (b) inclusive in m_X. No significant excesses (p-value < 0.01) over the background are found.

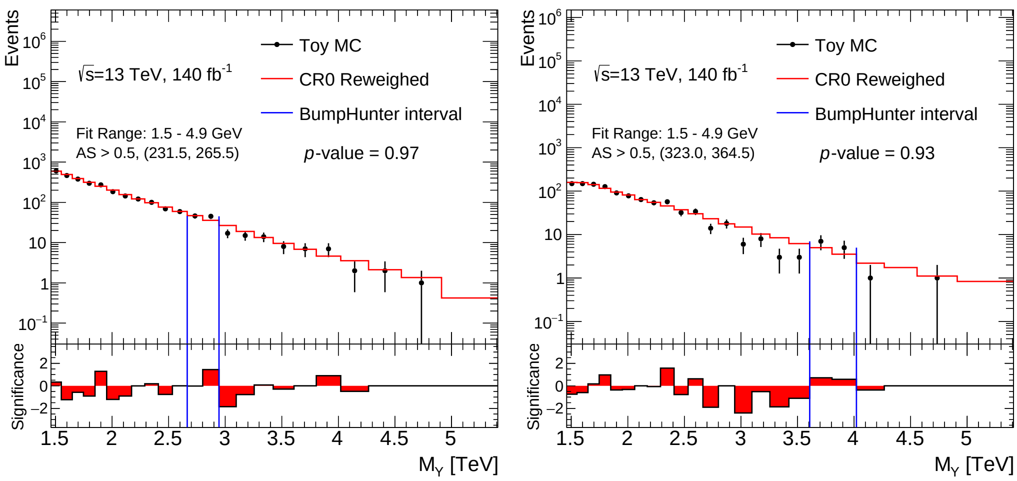


Figure 4.81: BUMPHUNTER spurious signal test plots with toys in sample bins of m_X (a) (231.5, 265.5) GeV and (b) (323.0, 364.5) GeV. No significant excesses over the background are found.

injections.

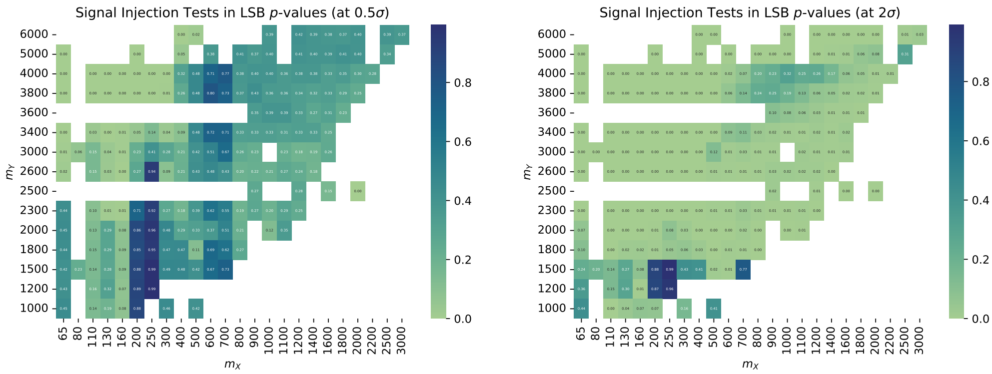
These tests reveal that the Anomaly region is capable of finding a 2-prong signal in a model-independent way. Primarily the score is sensitive to highly boosted

Table 4.6: BUMPHUNTER p-values for spurious signal with toys in m_X bins.

X Bin	p-value	BH Interval
(39.0, 92.0)	0.888	(1460.0, 1747.0)
(75.5, 95.5)	0.094	(2289.0, 2533.0)
(80.5, 101.0)	0.307	(3424.0, 4021.0)
(85.5, 107.0)	0.202	(4557.0, 5409.0)
(91.0, 112.5)	0.657	(2061.0, 2289.0)
(96.5, 118.5)	0.009	(1849.0, 2061.0)
(102.0, 124.5)	0.235	(2172.0, 2664.0)
(107.5, 130.5)	0.84	(1460.0, 1648.0)
(113.0, 137.0)	0.629	(1551.0, 1747.0)
(119.0, 143.5)	0.162	(3805.0, 4912.0)
(125.5, 150.0)	0.914	(2061.0, 2289.0)
(131.5, 156.5)	0.865	(4021.0, 4557.0)
(137.5, 163.5)	0.916	(1849.0, 2061.0)
(144.0, 170.5)	0.557	(3608.0, 4557.0)
(150.5, 177.5)	0.358	(3608.0, 6400.0)
(157.5, 184.5)	0.656	(3097.0, 3424.0)
(164.0, 192.0)	0.815	(1952.0, 2409.0)
(171.0, 199.5)	0.033	(1460.0, 1648.0)
(178.0, 207.0)	0.021	(3424.0, 3805.0)
(185.0, 215.0)	0.527	(1551.0, 2061.0)
(192.5, 223.0)	0.272	(3256.0, 3805.0)
(200.0, 231.0)	0.546	(1460.0, 1648.0)
(207.5, 239.5)	0.144	(1747.0, 2172.0)
(215.5, 248.0)	0.779	(1648.0, 1849.0)
(223.5, 256.5)	0.911	(2061.0, 2409.0)
(231.5, 265.5)	0.97	(2664.0, 2945.0)
(240.0, 274.5)	0.348	(4021.0, 5409.0)
(248.5, 283.5)	0.671	(2172.0, 2409.0)
(257.0, 293.0)	0.603	(2533.0, 3256.0)
⋮	⋮	⋮

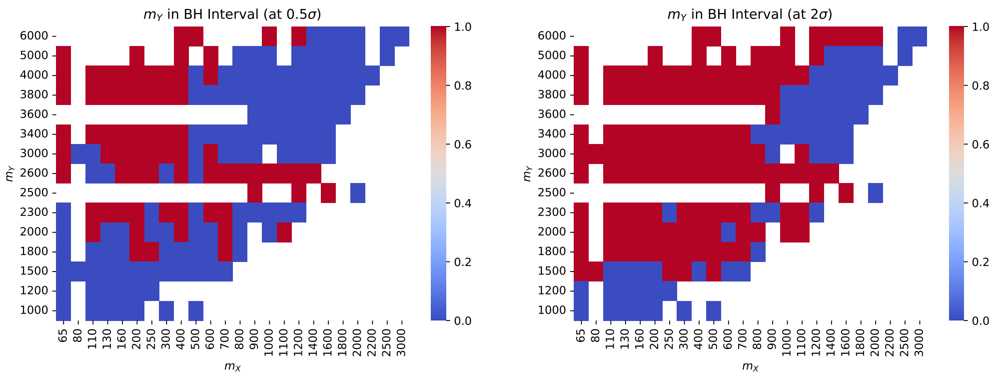
X Bin	p-value	BH Interval
⋮	⋮	⋮
(266.0, 302.5)	0.477	(4021.0, 4557.0)
(275.0, 312.5)	0.193	(2801.0, 4021.0)
(284.5, 322.5)	0.499	(4267.0, 4912.0)
(294.0, 332.5)	0.598	(3256.0, 3805.0)
(303.5, 343.0)	0.742	(2409.0, 2664.0)
(313.0, 353.5)	0.889	(1648.0, 1952.0)
(323.0, 364.5)	0.934	(3608.0, 4021.0)
(333.5, 375.5)	0.509	(2409.0, 2801.0)
(343.5, 387.0)	0.001	(2409.0, 2801.0)
(354.5, 398.5)	0.812	(4267.0, 4912.0)
(365.5, 410.0)	0.929	(1747.0, 1952.0)
(376.0, 422.0)	0.534	(1747.0, 2061.0)
(387.5, 434.0)	0.383	(4267.0, 4912.0)
(399.0, 446.5)	0.301	(1952.0, 2409.0)
(410.5, 459.0)	0.494	(2061.0, 2289.0)
(422.5, 472.0)	0.49	(2409.0, 2664.0)
(434.5, 485.5)	0.577	(2409.0, 3097.0)
(447.0, 499.0)	0.631	(2664.0, 2945.0)
(460.0, 512.5)	0.288	(1551.0, 1747.0)
(472.5, 526.5)	0.814	(2409.0, 2664.0)
(485.5, 541.0)	0.173	(1648.0, 2664.0)
(499.5, 555.5)	0.316	(2409.0, 2664.0)
(513.0, 570.5)	0.05	(3256.0, 3805.0)
(520.0, 592.5)	0.836	(1849.0, 2061.0)
(527.5, 621.5)	0.551	(1648.0, 1952.0)
(537.5, 659.0)	0.462	(2664.0, 2945.0)
(546.5, 711.0)	0.488	(1747.0, 2664.0)
(581.0, 3200.0)	0.018	(4021.0, 4557.0)

signals, where discovery potential drops off in the resolved area of the signal grid.



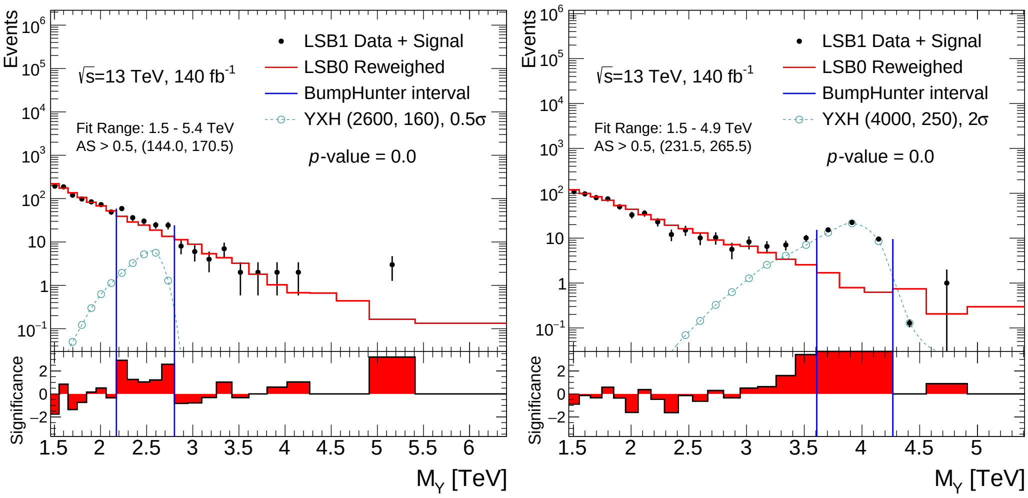
(a) (b)

Figure 4.82: BUMPHUNTER p-values for spurious signal and signal injection tests. (a) and (b) are respectively for 0.5σ and 2σ significance.



(a) (b)

Figure 4.83: Summary of signal injections where m_Y value is in BUMPHUNTER interval, for (a) 0.5σ and (b) 2σ significance. A 0 corresponds to the signal being outside the interval and 1 to it being inside.



(a) (b)

Figure 4.84: Signal injection BUMPHUNTER plots for the (m_Y, m_X) sample signal (a) (2600, 160) GeV at 0.5σ and (b) (4000, 250) GeV at 2σ significance.

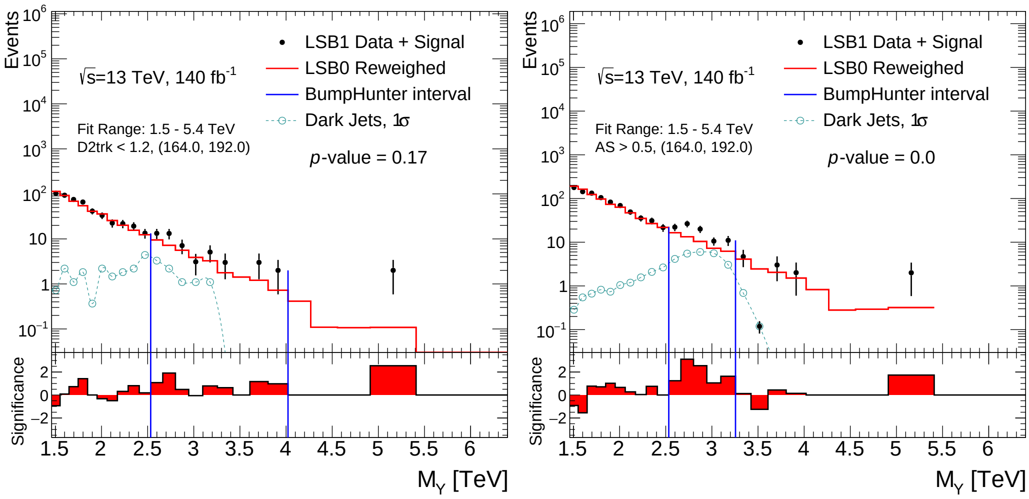
Comparison to D_2^{trk}

Because of the unsupervised training used in the anomaly region design, it is not expected to be as sensitive to the Two-prong $Y \to XH$ signals as the D_2^{trk} selection. However, its independence of specific signal features makes it more generalizable, and thus more sensitive to other kinds of jet topologies.

Figure 4.85 compares the sensitivity of the model-independent BumpHunter procedure after injecting the dark jet signal, defined in Section 4.2.2, with 1σ significance in the given m_X window. Two optimized selections are shown: the Two-prong merged selection, and the Anomaly model-independent selection. The AS preserves more of a resonant feature shape in m_{JJ}, resulting in a lower p-value for the background-only hypothesis and thus better sensitivity.

Tables 4.7 and 4.8 summarize the p-values for injecting the dark jet signal in each m_X bin both in the AS and D_2^{trk} regions.

The model-independent region produces p-values below threshold for various m_X bins where the D_2^{trk} region does not spot a significant excess, demonstrating that Anomaly region can be sensitive to a variety of signal substructure hypotheses of O(1) σ significance in the SR.



(a) (b)

Figure 4.85: BumpHunter signal injection tests with the dark jet signal, comparing the D_2^{trk} exclusion selection (a) with the AS discovery selection (b), in the m_X bin [164.0,19.20] GeV. The AS selection gives more of a resonant and therefore observable feature in these unusual signals.

Table 4.7: AS region BUMPHUNTER p-values for the dark jet signals injection tests in m_X bins.

X Bin	p-value	BH Interval	X Bin	p-value	BH Interval
(39.0, 92.0)	0.003	(2533.0, 4021.0)	⋮	⋮	⋮
(75.5, 95.5)	0.001	(2172.0, 2801.0)	(266.0, 302.5)	0.08	(2172.0, 3256.0)
(80.5, 101.0)	0	(2172.0, 2801.0)	(275.0, 312.5)	0.289	(2289.0, 3256.0)
(85.5, 107.0)	0	(2289.0, 2801.0)	(284.5, 322.5)	0.057	(2289.0, 3256.0)
(91.0, 112.5)	0.059	(2664.0, 2945.0)	(294.0, 332.5)	0.005	(2289.0, 3256.0)
(96.5, 118.5)	0.041	(2533.0, 4021.0)	(303.5, 343.0)	0.006	(2289.0, 3424.0)
(102.0, 124.5)	0.071	(2533.0, 4021.0)	(313.0, 353.5)	0.006	(2289.0, 3805.0)
(107.5, 130.5)	0.038	(2533.0, 4021.0)	(323.0, 364.5)	0.021	(2289.0, 3805.0)
(113.0, 137.0)	0	(2172.0, 3608.0)	(333.5, 375.5)	0.155	(2289.0, 3608.0)
(119.0, 143.5)	0	(2172.0, 3608.0)	(343.5, 387.0)	0.159	(2061.0, 3097.0)
(125.5, 150.0)	0	(2172.0, 3608.0)	(354.5, 398.5)	0.178	(2061.0, 3097.0)
(131.5, 156.5)	0	(2172.0, 3608.0)	(365.5, 410.0)	0.289	(2533.0, 2801.0)
(137.5, 163.5)	0	(2172.0, 3608.0)	(376.0, 422.0)	0.057	(2061.0, 3097.0)
(144.0, 170.5)	0.001	(2172.0, 3424.0)	(387.5, 434.0)	0.173	(2664.0, 3097.0)
(150.5, 177.5)	0	(2533.0, 3424.0)	(399.0, 446.5)	0.281	(2409.0, 3097.0)
(157.5, 184.5)	0	(2533.0, 3424.0)	(410.5, 459.0)	0.091	(2664.0, 3608.0)
(164.0, 192.0)	0	(2533.0, 3256.0)	(422.5, 472.0)	0.616	(1460.0, 1648.0)
(171.0, 199.5)	0.001	(2533.0, 3256.0)	(434.5, 485.5)	0.618	(1460.0, 1747.0)
(178.0, 207.0)	0.035	(2409.0, 3256.0)	(447.0, 499.0)	0.065	(1460.0, 1747.0)
(185.0, 215.0)	0.041	(2409.0, 3256.0)	(460.0, 512.5)	0.154	(1460.0, 1747.0)
(192.5, 223.0)	0.034	(2409.0, 3256.0)	(472.5, 526.5)	0.474	(1551.0, 1747.0)
(200.0, 231.0)	0.078	(2533.0, 4021.0)	(485.5, 541.0)	0.352	(2172.0, 2664.0)
(207.5, 239.5)	0.009	(2801.0, 3256.0)	(499.5, 555.5)	0.159	(2172.0, 2801.0)
(215.5, 248.0)	0.073	(2801.0, 3256.0)	(513.0, 570.5)	0.102	(2172.0, 2409.0)
(223.5, 256.5)	0.402	(2533.0, 4021.0)	(520.0, 592.5)	0.366	(3424.0, 4267.0)
(231.5, 265.5)	0.217	(2664.0, 3256.0)	(527.5, 621.5)	0.201	(2289.0, 3097.0)
(240.0, 274.5)	0.047	(2664.0, 3097.0)	(537.5, 659.0)	0.304	(2289.0, 3097.0)
(248.5, 283.5)	0.146	(2664.0, 4267.0)	(546.5, 711.0)	0.47	(2289.0, 3256.0)
(257.0, 293.0)	0.297	(2945.0, 4267.0)	(581.0, 3200.0)	0.246	(2409.0, 3256.0)
⋮	⋮	⋮			

Table 4.8: D_2^{trk} region BUMPHUNTER p-values for the dark jet signals injection tests in m_X bins.

X Bin	p-value	BH Interval	X Bin	p-value	BH Interval
(39.0, 92.0)	0.049	(2533.0, 3097.0)	⋮	⋮	⋮
(75.5, 95.5)	0.083	(2533.0, 3256.0)	(266.0, 302.5)	0.217	(2172.0, 3097.0)
(80.5, 101.0)	0	(2533.0, 3424.0)	(275.0, 312.5)	0.233	(2172.0, 3097.0)
(85.5, 107.0)	0.044	(2289.0, 3424.0)	(284.5, 322.5)	0.268	(2172.0, 3256.0)
(91.0, 112.5)	0.05	(2533.0, 2945.0)	(294.0, 332.5)	0.01	(2289.0, 3097.0)
(96.5, 118.5)	0.024	(1849.0, 3256.0)	(303.5, 343.0)	0.045	(2289.0, 3424.0)
(102.0, 124.5)	0.041	(2533.0, 3256.0)	(313.0, 353.5)	0.004	(2289.0, 3805.0)
(107.5, 130.5)	0.103	(2801.0, 3256.0)	(323.0, 364.5)	0.017	(2289.0, 3608.0)
(113.0, 137.0)	0.094	(2801.0, 3256.0)	(333.5, 375.5)	0.032	(2289.0, 2533.0)
(119.0, 143.5)	0.115	(2409.0, 3097.0)	(343.5, 387.0)	0.008	(2289.0, 2801.0)
(125.5, 150.0)	0.074	(2172.0, 2945.0)	(354.5, 398.5)	0.025	(2289.0, 2801.0)
(131.5, 156.5)	0.023	(2061.0, 3256.0)	(365.5, 410.0)	0.014	(2289.0, 2801.0)
(137.5, 163.5)	0.023	(2172.0, 3424.0)	(376.0, 422.0)	0.029	(2061.0, 3256.0)
(144.0, 170.5)	0.036	(2533.0, 4021.0)	(387.5, 434.0)	0.176	(2664.0, 4021.0)
(150.5, 177.5)	0.007	(2533.0, 3424.0)	(399.0, 446.5)	0.302	(2289.0, 3805.0)
(157.5, 184.5)	0.019	(2533.0, 2801.0)	(410.5, 459.0)	0.063	(2409.0, 3805.0)
(164.0, 192.0)	0.174	(2533.0, 4021.0)	(422.5, 472.0)	0.352	(2409.0, 2801.0)
(171.0, 199.5)	0.025	(2533.0, 3424.0)	(434.5, 485.5)	0.238	(1952.0, 2801.0)
(178.0, 207.0)	0.102	(2061.0, 2945.0)	(447.0, 499.0)	0.263	(1551.0, 1747.0)
(185.0, 215.0)	0.004	(2061.0, 3256.0)	(460.0, 512.5)	0.074	(1551.0, 1747.0)
(192.5, 223.0)	0	(2061.0, 3805.0)	(472.5, 526.5)	0.615	(1648.0, 2533.0)
(200.0, 231.0)	0.01	(2061.0, 3805.0)	(485.5, 541.0)	0.677	(2061.0, 2533.0)
(207.5, 239.5)	0.008	(2061.0, 3256.0)	(499.5, 555.5)	0.373	(2289.0, 2533.0)
(215.5, 248.0)	0.014	(2289.0, 2945.0)	(513.0, 570.5)	0.41	(2289.0, 2533.0)
(223.5, 256.5)	0.026	(2289.0, 3256.0)	(520.0, 592.5)	0.495	(3805.0, 4912.0)
(231.5, 265.5)	0.034	(2409.0, 3256.0)	(527.5, 621.5)	0.267	(2289.0, 2945.0)
(240.0, 274.5)	0.032	(2409.0, 3097.0)	(537.5, 659.0)	0.292	(2289.0, 2945.0)
(248.5, 283.5)	0.16	(2409.0, 3805.0)	(546.5, 711.0)	0.422	(2289.0, 2664.0)
(257.0, 293.0)	0.167	(2172.0, 2801.0)	(581.0, 3200.0)	0.514	(2409.0, 2664.0)
⋮	⋮	⋮			

4.7.3 Two-prong Regions

The background-only and signal plus background fits are implemented using TREXFITTER[5] framework. It is performed on m_{JJ} histograms in each of the defined $m_X$bins.
A simultaneous binned likelihood fit is performed across the Two-prong merged and resolved regions. The normalization of the data-driven background estimation is allowed to float in all control, validation, and signal regions for both merged and resolved selections, with each normalization factor being fit independently as the m_X categories are overlapping.
The parameter of interest in the statistical analysis is the signal strength μ, defined as a scale factor on the total number of signal events with respect to the nominal predicted by a σ=1 pb assumption. The background only hypothesis corresponds to μ = 0, and the hypothesis of the full signal plus background gives μ = 1.
A test statistic based on the profile likelihood ratio using the lowest order asymptotic approximation is used to test the models proposed by the signal grid.
Systematic uncertainties are incorporated into the fit as nuisance parameters (NPs) with Gaussian constraints. Both the signal strength and all signal systematic NPs are correlated across merged and resolved regions. The significance of an excess observed in data over the background prediction is quantified by the local p_0, which is the probability of the background only model to produce an excess at least as large as the one observed.
Upper limits on the signal cross section are set using a modified frequentist method (CL_s) [104]. The values CL_{s+b} and CL_b are computed using an asymptotic method [71] and correspond to the p-values for the signal-plus-background and background-only hypothesis, respectively.

Background-only fits

Figure 4.86 shows the pre- and post-fit m_{JJ} distributions, for the m_X bin in the range of 284.5 GeV to 322.5 GeV, to illustrate the fit functionality in a blinded data region (LSB VR).
Figure 4.87 shows the obtained normalization factors from the LSB VR and the pulls of the nuisance parameters, which reflect the background systematic uncertainties on the merged and resolved regions separately.

[5] https://trexfitter-docs.web.cern.ch/

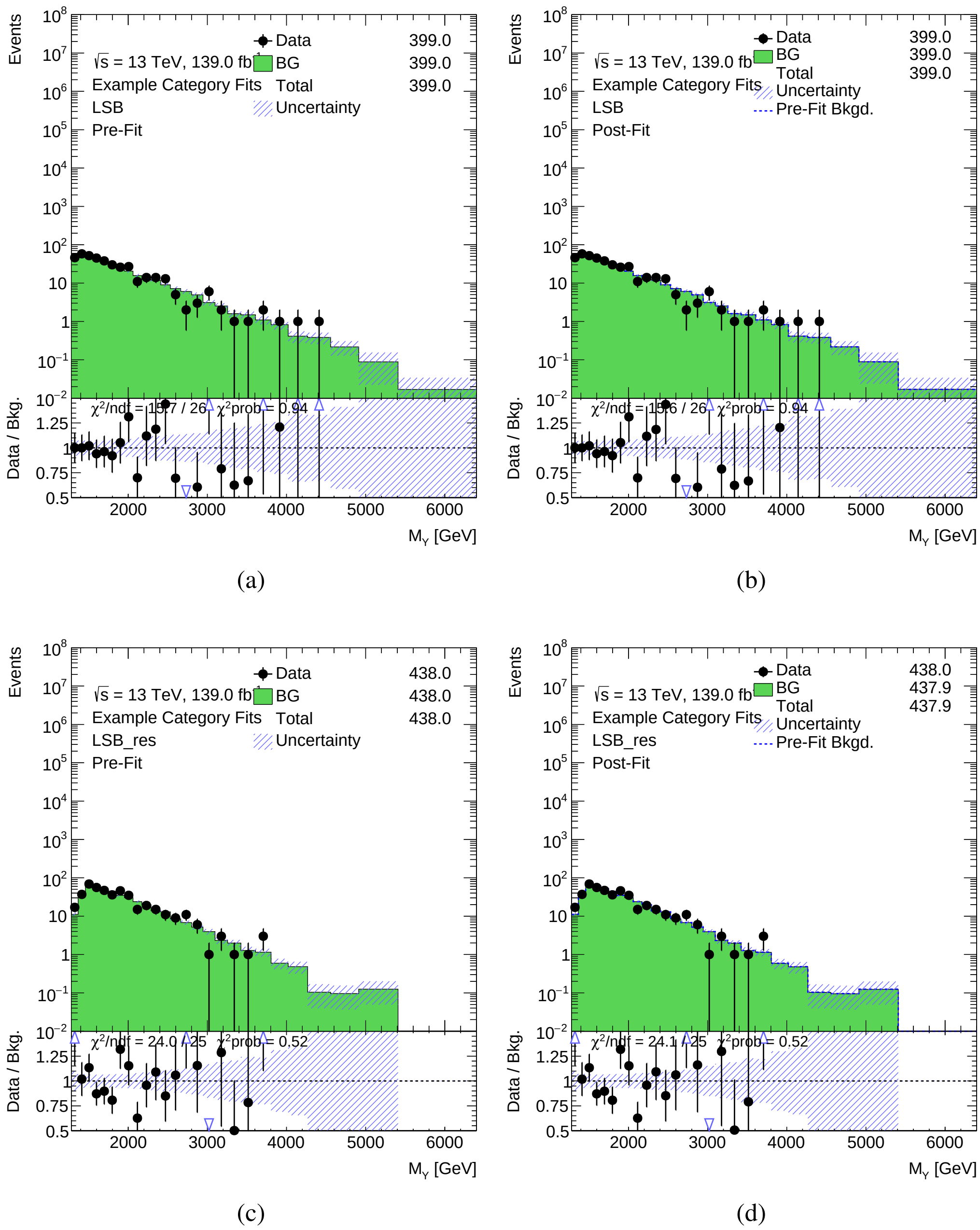
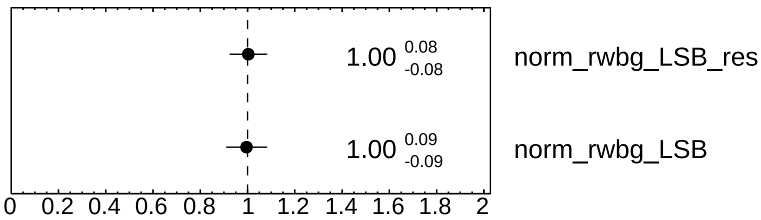
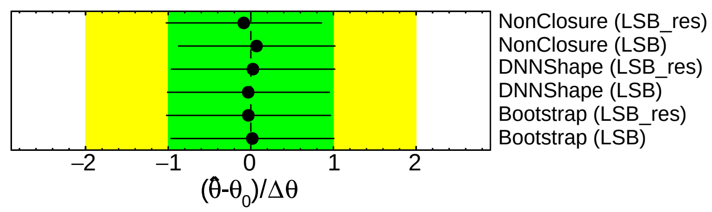


Figure 4.86: Distribution of m_{JJ} in the LSB VR within the m_X window [284.5, 322.5] GeV, pre (left) and post (right) fit with systematics, separately for the Two-prong merged (top) and resolved (bottom) selection criteria



(a)



(b)

Figure 4.87: Normalization and nuisance parameter pulls in the VR background-only fit within the m_X window [284.5, 322.5].

Signal plus background fits

The LSB VR is fit with a signal-plus-background hypothesis to test the health of the fit when signal systematic NPs are considered.
The result of the combined merged+resolved fit for (m_Y= 4000, m_X= 300) GeV can be seen in Figure 4.88. Here the signal has an arbitrary 1fb normalization and does not represent a realistic expectation in the LSB, but is useful to observe fit behavior. A ranking plot of all uncertainties for (m_Y= 4000, m_X= 300) GeV is also provided. No significant pulls or profiles are observed.

Signal plus background fits with Asimov data in the SR are also studied to build an expected sensitivity.
Each S+B fit of m_Y in a bin of m_X produces a CLs value for the signal point that is enriched by that particular (m_Y,m_X) bin. This can be translated to an upper limit on the $Y \to XH$ signal cross section.
An example 1D limit can be seen in Figure 4.7.3 for the (m_Y=4000,m_X=300) GeV and [284.5, 322.5] GeV m_X bin shown throughout this section.

Spurious signal tests

Extensive studies were performed on the spurious signal in the Two-prong region, specifically on whether a dedicated spurious signal uncertainty was necessary.
The conclusion from these efforts is that the Non-closure uncertainty is sufficient to cover any fittable spurious signal in data and the study is documented below.

The spurious signal tests is done via S+B fits, and the fitted spurious signal is measured in σ significance as the fitted μ divided by its uncertainty. In each fit, μ is restricted to be positive.
First, Asimov data are drawn from the background estimate in the SR, with statistics equivalent to what is expected in the unblinded data. 100 toys are thrown to emulate an imperfect distribution in data, with statistical fluctuations that may be locally incompatible with the background prediction.
Figure 4.91 shows the result of an example signal fit (m_Y= 2600 GeV, m_X=300 GeV) in this configuration, including the pre- and post-fit m_{JJ} distribution. Also shown is the histogram of fitted μ values for all pseudo-experiments in this fit, which takes the shape of a Gaussian, indicating that no systematic spurious signal is evident.
Figure 4.92 shows the fitted spurious signal significance in this Asimov dataset across the entire $Y \to XH$ signal grid. This result provides a sanity check that the

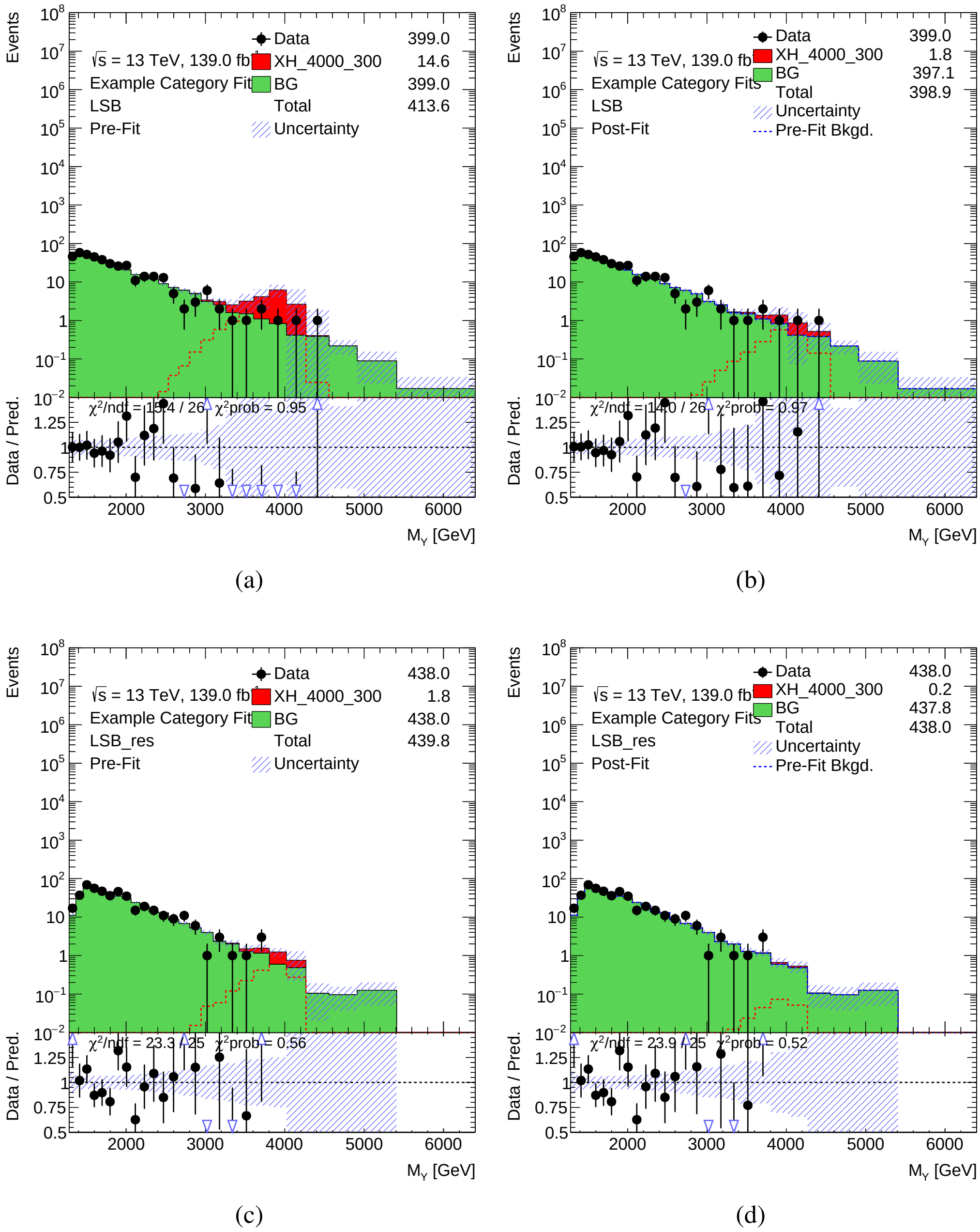


Figure 4.88: Pre- and post-fit m_{JJ} distributions in the Two-prong resolved (top) and merged (bottom) selections for a combined merged+resolved fit in the LSB for signal point (m_Y= 4000, m_X= 300) GeV in the m_X window [284.5, 322.5] GeV.

fit framework is robust and no deviations are found in a toy Asimov dataset that is

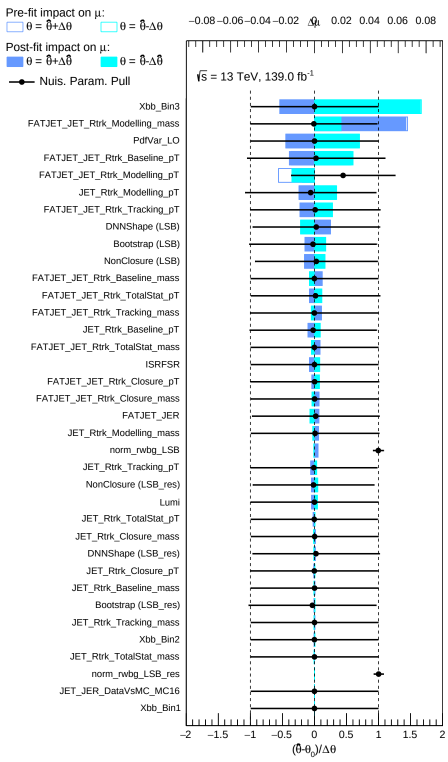


Figure 4.89: A ranking plot of all signal plus background systematics in the combined merged+resolved selection LSB for signal point (m_Y= 4000, m_X= 300) GeV in the m_X window [284.5, 322.5] GeV.

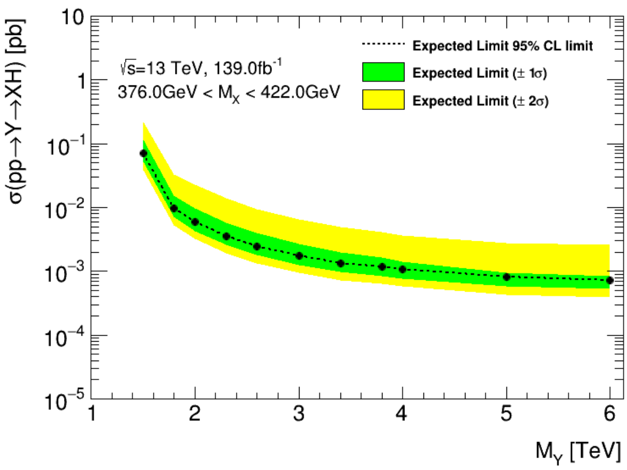
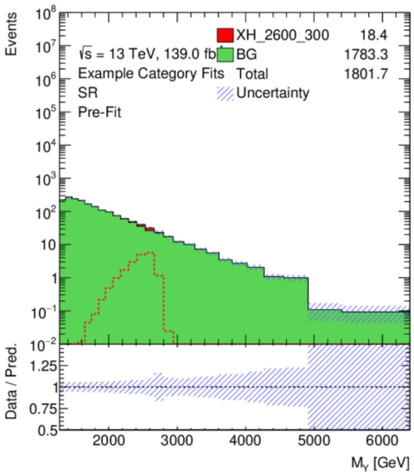
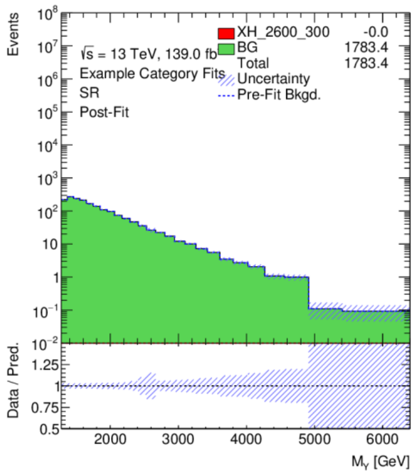


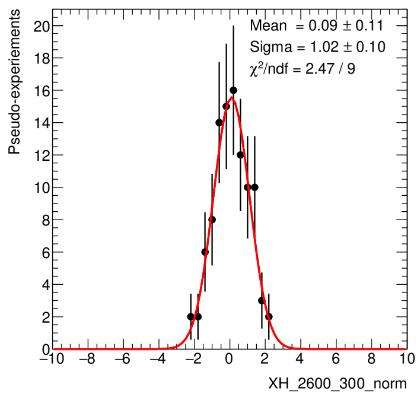
Figure 4.90: 1D expected cross section limit vs. m_Y in the m_X window [284.5, 322.5] GeV, with Asimov data in the SR.



(a)



(b)



(c)

Figure 4.91: Example spurious signal S+B fit to Asimov data in the SR, for point (m_Y= 2600 GeV, m_X=300 GeV). (a) and (b) are respectively pre- anf post-fit distributions, while (c) is the histogram of fitted μ values for all pseudo-experiments, which takes the shape of a Gaussian, indicating that no systematic spurious signal is evident.

directly drawn from the background estimate.

These checks were done to real data in the VR as well, which provides a more realistic scenario in which spurious signal could appear, as real statistical fluctuations are present. Figure 4.93 shows the fitted spurious signal significance in the validation region across the entire $Y \to XH$ signal grid. No significant excesses are found, nor trends or correlations to any particular region of the grid, which may have indicated systematic effects not covered by the included systematics.

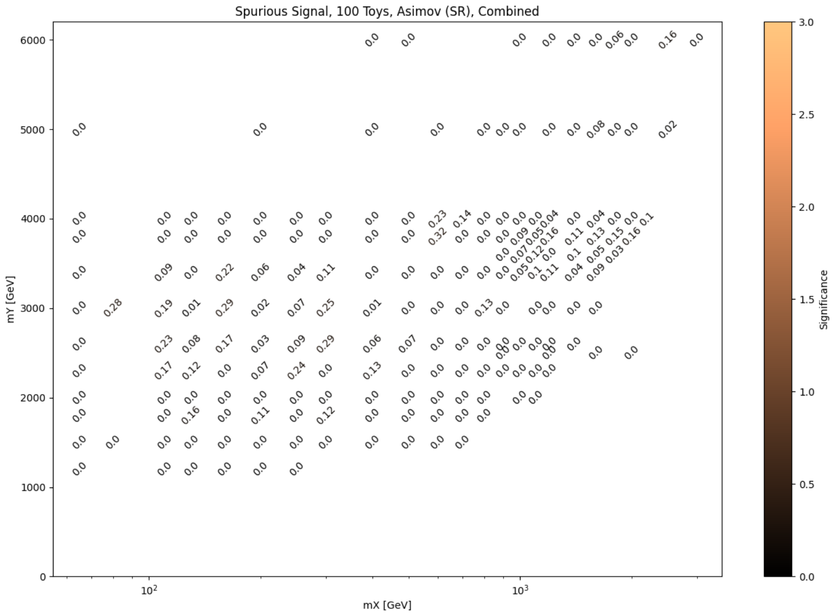


Figure 4.92: Fitted spurious signal excess in σ for an S+B fit to Asimov data in the SR.

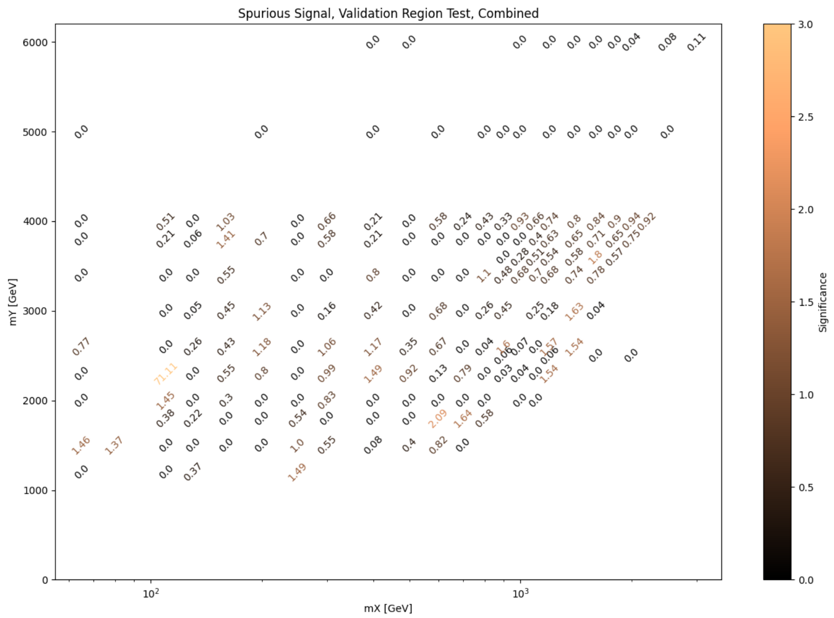


Figure 4.93: Fitted spurious signal excess in σ for an S+B fit to data in the VR.

4.8 Results

This section shows obtained results from unblinded fits in the two kind of signal regions of this analysis.

4.8.1 Anomaly region

Results of background-only fits of the m_{JJ} distribution across all m_X categories in the anomaly SR show good compatibility of data to expected background, after incorporating all statistical and systematic uncertainties.
Figure 4.94 shows the distribution of p_0 values across both m_Y and m_X, where the background is the result of a background-only fit performed with all statistical and systematic uncertainties.

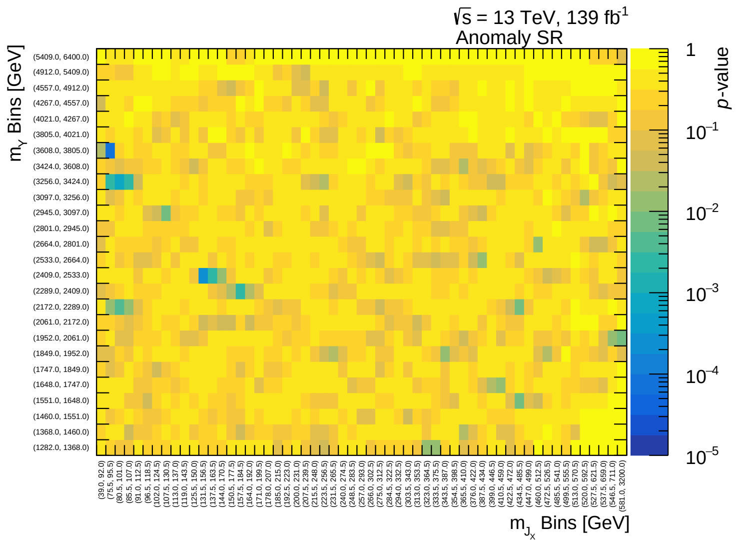


Figure 4.94: The distribution of observed p_0 values across all m_Y and m_X bins in the anomaly signal region, comparing data to the background estimation generated by a background-only fit. All statistical and background systematic uncertainties are considered in the p_0 calculation. The lowest observed p_0 corresponds to the bin with $m_Y$within [3608, 3805] GeV and $m_X$within [75.5, 95.5] GeV.

The largest observed excess in the anomaly SR is in the m_X window [75.5, 95.5] GeV, where the BUMPHUNTER interval covers the m_Y range between bin edges of 3424 and 3805 GeV. The excess corresponds to a p-value of 8.8×10^{-3}, where the BUMPHUNTER considers only statistical uncertainties. The m_{JJ} distribution corresponding to this window is shown in Figure 4.95, along with the post-fit expected background. Studies of the individual jet mass distributions in this m_X category do not reveal any excesses in data that can be consistent with a resonant particle decay (as shown in Appendix B.3).

Given the number of individual search regions in this analysis, the impact of the trials factor is significant. A calculation is made to determine the global significance of this deviation accounting for all the overlapping m_X bins, where the overlapping bin edges are used to define exclusive, non-overlapping bins in m_X; an integer is drawn from a Poisson distribution with a mean of the background expectation in each exclusive (m_Y, m_X) bin; this yield is summed across exclusive bins to create a toy estimate for each overlapping bin in which the p-value is computed; and this procedure is repeated N times where N is the number of events inclusively across all exclusive, non-overlapping bins. The trials factor is then the fraction of such pseudo-experiments where the largest observed excess is greater than that observed in the single exclusive (m_Y, m_X) bin.

This calculation yields a global significance of 1.43σ for this excess.

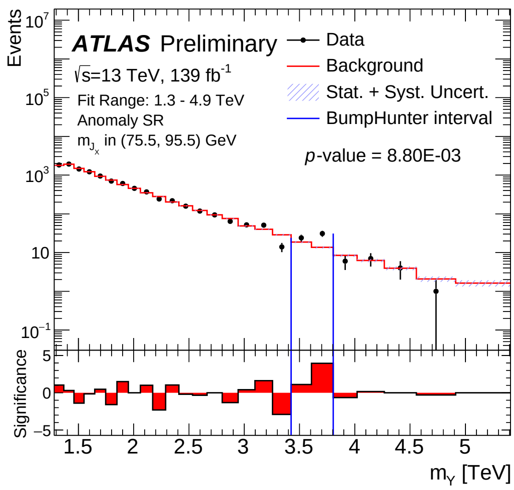
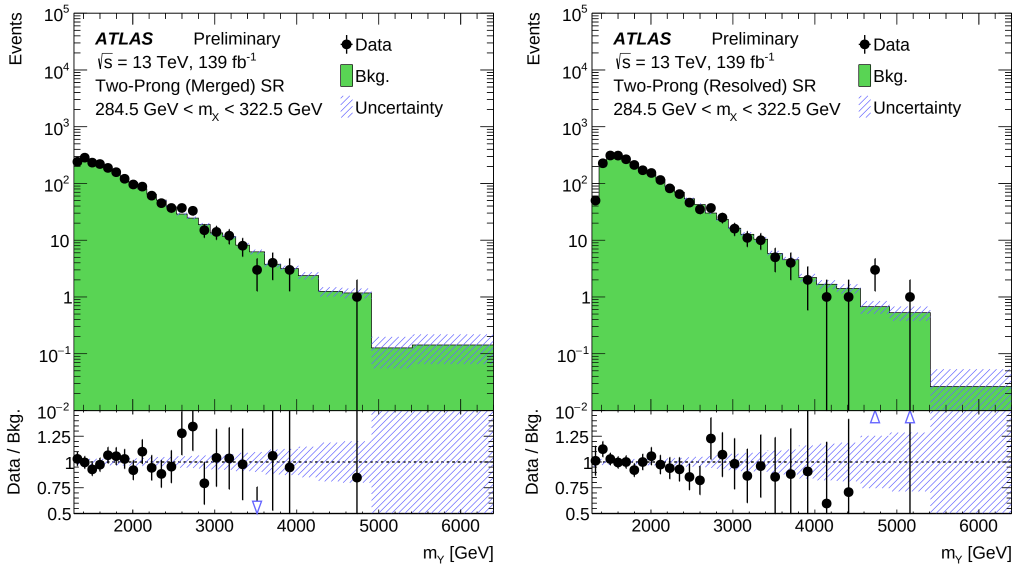


Figure 4.95: The m_{JJ} distribution associated to the m_X bin [75.5, 95.5] GeV in the Anomaly SR, where the background is determined by a background-only fit to data with both statistical and systematic uncertainties included. The p-value of 8.8×10^{-3} pertains to the interval of m_Y bins chosen by `BumpHunter`, shown in blue, to have the most significant excess over expected background. The p-value and the lower panel of per-bin significances are computed with statistical uncertainties only. Incorporating the trials factor for all m_Y and overlapping m_X bins in the search, the single bin with the most significant excess has a global significance of 1.43σ. [7]

4.8.2 Two-prong regions

Results for the two-prong SRs are similarly derived by performing background-only fits and scanning with `BumpHunter` for incompatibility with data.
No significant deviations of data are observed with respect to the predicted background beyond expected statistical fluctuations, in either the merged or resolved SR.

An example m_{JJ} distribution in both the merged and resolved SRs for the m_X bin [284.5, 322.5] GeV is shown in Figure 4.96, along with the background estimation that is determined from a background-only fit accounting for all uncertainties. Figure 4.97 shows a summary of the per-bin p-values in each m_Y bin for selected



(a) (b)

Figure 4.96: Reconstructed m_Y distributions of background and data in the SR, for the merged (a) and resolved (b) selections, in the m_X bin [284.5, 322.5] GeV. The background is determined by a background-only fit to the data with all statistical and systematic uncertainties included. The ratio of the observed data to the background is shown in the lower panel. The uncertainty band includes both statistical and systematic effects. [7]

m_X bins that are centered on key X candidate mass hypotheses, namely at the W, Z, and H boson masses.

Given the absence of significant excesses in the data, signal-plus-background fits are performed to determine the 95% CL upper limit on the cross section of the

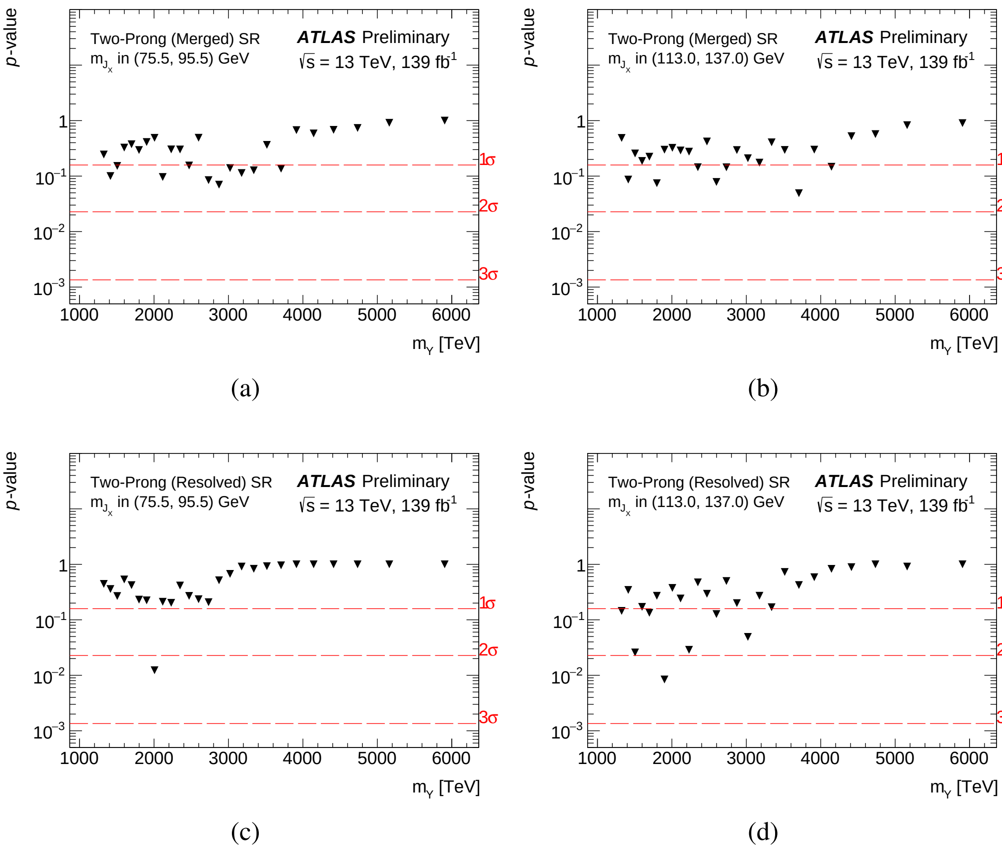


Figure 4.97: The p-value per m_Y bin for both two-prong SRs, calculated using all systematic and statistical uncertainties on the background estimation. Two m_X bins are shown, [75.5, 95.5] and [113.0, 137.0] GeV, which corresponds to a window containing the W/Z and Higgs boson mass respectively. Events are thus split into merged W/Z window (a), merged Higgs window (b), resolved W/Z window (c), and resolved Higgs window (d). The background is determined by a background-only fit to the data with all statistical and systematic uncertainties included. In both m_X windows, the p-value approximates a constant value of 0.5 for the high Y mass region of the resolved SR, as this region of phase space is far more likely to produce a highly boosted J_X that falls in the merged SR selection. [7]

$Y \to XH$ process.
A summary of the expected and observed limits in the 2D grid of the $Y \to XH$ signals is given in Figure 4.98, combining results from the Two-prong merged and resolved regions. The analysis is most sensitive in the very boosted regime, where the Y mass is approximately an order of magnitude larger than the X mass. Sensitivity is worst in the highly resolved regime, due to the required large-R J reconstruction of the Higgs boson which sculpts signal efficiency to high momentum X particles.
The observed limits range from cross sections of 0.341 fb for the signal point (m_Y= 5000 GeV, m_X= 600 GeV), to 1.22 pb for the signal point (m_Y= 2500 GeV, m_X= 2000 GeV).

The data in the Anomaly and Two-prong SRs can be used to provide a benchmark comparison of sensitivity across the set of large-R jet decays considered for the X, thereby providing a metric for assessing the level of signal model-dependence in both regions. 95% CL upper limits on the production cross section of several benchmark signals are calculated for all three signal region selections, by injecting signal into the data until the BumpHunter p-value exceeds a significance of 2σ. Seven signal points are considered in this study, including three highly boosted $Y \to XH$ points, one resolved $Y \to XH$ point, and the three alternate jet topologies.
As systematic uncertainties on the signal efficiency of the anomaly score are not assessed, this comparison is performed using only statistical uncertainties and a post-fit background estimation in the limit calculation. Since the merged region uses D_2^{trk} and thus explicitly tags on the two-prong substructure of the X in the generated $Y \to XH$ grid, it is possible that these regions will outperform the fully unsupervised approach on the $Y \to XH$ signals.
Figure 4.99 provides a summary of the obtained limits from the signal injection tests, comparing all three signal region selections for the seven benchmark signal points. For points where the X is highly boosted and thus the anomaly score is most sensitive, the upper limit on the cross section is approximately the same across the merged and anomaly SRs. The signal model-independent aspect of the anomaly detection used in the anomaly SR is evident through improved limits on the alternative substructure signals.
Notably, for the dark jets the limit is nearly an order of magnitude lower than that provided by the two-prong regions, which underlines the strength of the model-independent approach particularly for signals that are challenging to characterize with existing high-level variables.

While these results cover areas of phase space that have not been previously studied directly by other searches, some analysis selections are highly correlated to those of

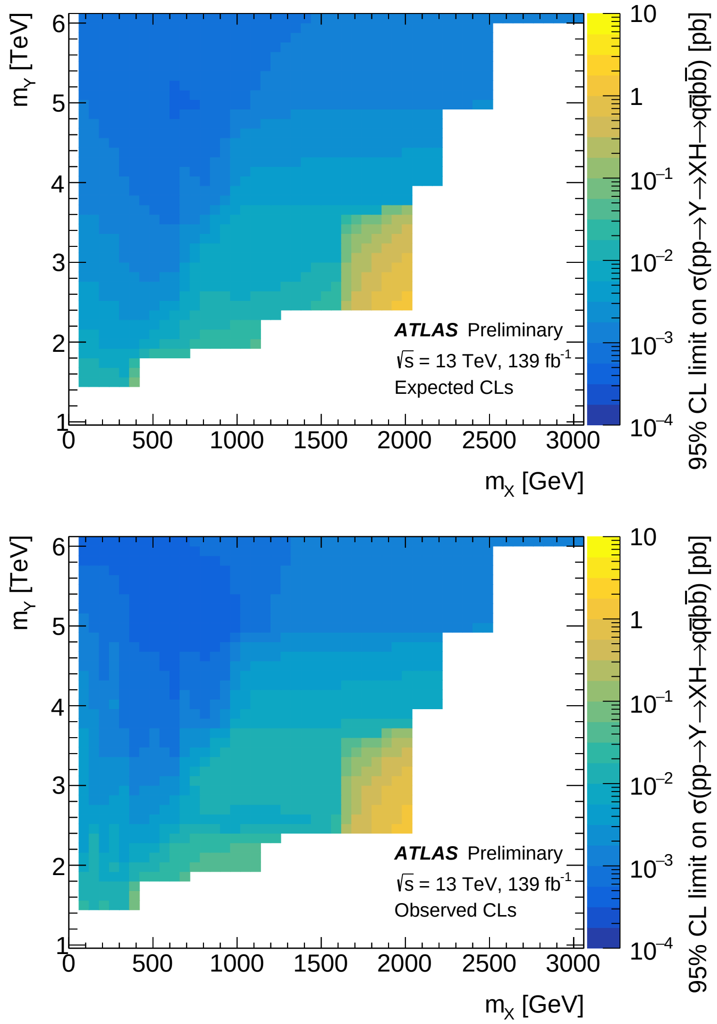


Figure 4.98: The expected (a) and observed (b) 95% CL limits on the cross-section $\sigma(pp \to Y \to XH \to q\bar{q}b\bar{b})$ in pb in the two-dimensional space of m_Y versus m_X, obtained from a simultaneous fit of both merged and resolved two-prong signal regions with all statistical and systematic uncertainties. A bilinear interpolation procedure is applied to provide results in between fully simulated signal points. The observed limits range from 0.341 fb for the signal point (m_Y= 5000 GeV, m_X= 600 GeV) to 1.22 pb for the signal point (m_Y= 2500 GeV, m_X= 2000 GeV). The boundaries of the limit are defined by the simulated signal grid. [7]

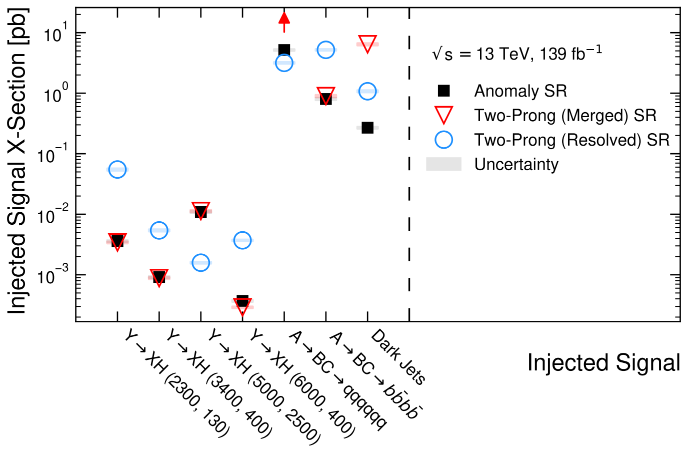


Figure 4.99: The 95% CL upper limit on the cross section of seven benchmark signal processes, comparing the anomaly, two-prong merged, and two-prong resolved signal region selections. The cross section is obtained by injecting signal into the observed data until a 2σ sensitivity is achieved as determined by the BumpHunter interval p-value. The background estimation is the result of a fit with all background uncertainties are considered, but no signal systematic uncertainties are used. The anomaly SR provides competitive sensitivity to all signals, while the two-prong regions fail to give good limits for at least one of the signal points considered.

other recent ATLAS dijet resonance searches. The m_X bin of [75.5, 95.5] would be sensitive to the VH resonance hadronic final state, which is covered by a dedicated analysis using the same dataset [37]. The approach here differs in both vector boson tagging and Higgs boson tagging approaches, but provides a similar 95% CL upper limit on the production cross section of a 3 TeV resonance. Due to its generality, the anomaly SR is expected to be sensitive to the same signatures as the weakly supervised dijet resonance search [36], though a direct comparison is not provided due to the assumptions made here of the Higgs boson mass and decay.

Conclusions

LHC is the perfect test-bench for exploring new Physics scenarios. The energies reached at the protons collider allow to directly probe the production of new heavy resonances. This thesis work is inserted in such a research context, in particular it is about a search for a new TeV-scale narrow-width boson Y, decaying to a Standard Model Higgs boson H and a new particle X with a mass O(10-1000) GeV. It was performed with 139 fb^{-1}of pp collisions collected by the ATLAS detector during Run-2.

The analysis focuses on a fully hadronic final state, where the X and H bosons are boosted such that their daughter particles are collimated.
Higgs decay is tagged with the *Hbb* tagger, a neural net-based classifier which separates boosted Higgs bosons decaying to $b\bar{b}$ from top quark and QCD jets. Different tagging strategies are applied on X jet: a fully unsupervised Variational Recurrent Neural Network is trained over jets in data to define the Anomaly Signal Region (SR), which selects the X particle solely based on its sub-structural incompatibility with background jets; in parallel, two supplementary SRs are designed to separately reconstruct merged and resolved decays of a nominal two-prong X benchmark signal. The latter has been inserted to recover sensitivity in an unexplored region of the phase space, where the large-R jet is not sufficiently efficient in reconstructing the boosted X decay.

Another important improvement introduced, with respect to a previous round of the analysis performed with 36 fb^{-1} [32], is the usage of a Deep Neural Network (DNN) for the data-driven background estimation. The background to the fully hadronic final state is composed primarily of jets from QCD processes, for which simulations are known to have limited precision, therefore it has been directly obtained from data in a control region and reweighting factors to simulate the SR kinematics have been derived by estimating the likelihood ratio between the SR and the control region, with the help of a DNN.
Only another ATLAS analysis [12] adopted it so far for the background estimation, being this another reason to consider this analysis at the forefront of the techniques

used.

In both the Anomaly and the Two-prong SRs no significant deviations have been observed in data with respect to the background-only hypothesis. The largest excess is found in the anomaly SR with a global significance of 1.43 σ, considering all m_X and m_Y bins.
The strength of the model-independent approach in the Anomaly region is evident particularly for signals that are challenging to characterize with existing high-level variables.
Results in the Two-prong SRs are interpreted as upper limits at 95% confidence level on the cross section $\sigma(pp \to Y \to XH \to q\bar{q}b\bar{b})$, across the two-dimensional space where m_Y is between 1.5 and 6 TeV and m_X is between 65 and 3000 GeV. The lowest limit of 0.341 fb is achieved in the merged regime for the signal point (m_Y= 5000 GeV, m_X= 600 GeV), while the highest of 1.22 pb is achieved for the highly resolved point (m_Y= 2500 GeV, m_X= 2000 GeV).

Appendices

APPENDIX A

ATLAS New Small Wheels

In order to benefit from the expected high luminosity performance provided by the Phase-I upgraded LHC, the first stations of the ATLAS muon end-cap system (Small Wheels) need to be replaced to operate in a high background radiation region (up to $15\,\mathrm{kHz/cm^2}$) while reconstructing muon tracks with high precision as well as furnishing information for the Level-1 trigger. Figure A.1 shows a cross section of the ATLAS detector in z-y plane. The barrel system covers the region of $|\eta| < 1.0$ whereas the end-cap system covers $1.0 < |\eta| < 2.7$ for muon tracking and $1.0 < |\eta| < 2.4$ for Level-1 trigger.
The Phase-I upgrade of the ATLAS muon spectrometer focuses on the end-cap region, because at high luminosity the following two points are of particular importance:

- The performance of the muon tracking chambers degrades with the expected increase of background rate. Precise measurement of the muon p_T requires the presence of track segments in all three muon stations. Losing Small Wheel segments leads to a loss of high quality muon tracking. MDTs lose hits at high occupancy conditions due to the long dead time (of about $800\,\mathrm{ns}$). The CSCs, located at the region with the highest background conditions and having only four detection layers, are also affected by high occupancy arising from overlaps of multiple hits on the readout strips. This degradation is detrimental for the performance of the ATLAS detector.

- The Level-1 muon trigger in the end-cap region is based on track segments in the TGC chambers of the middle muon station located after the end-cap toroid magnet. The transverse momentum of the muon is determined by the angle of the segment with respect to the direction pointing to the interaction point. A significant part of the muon trigger rate in the end-caps is background. Low energy particles generated in the material located between the Small Wheel and the middle muon station, generally protons, produce fake triggers by hitting the end-cap trigger chambers at an angle similar to that of real high p_T muons. An analysis of 2012 data demonstrates that approximately 90% of the muon triggers in the end-caps are fake. As a consequence, the rate of the Level-1 muon trigger in the end-cap is eight to nine times higher than that in the barrel region.

Both of these two issues represent a serious limitation on the ATLAS performance beyond design luminosity: reduced acceptance of good muon tracking and an unacceptable rate of fake high p_T Level-1 muon triggers coming from the forward direction. In order to solve the two problems together, ATLAS proposes to replace the present Small Wheels with the New Small Wheels (NSWs). The NSW is a set of precision tracking and trigger detectors able to work at high rates with excellent real-time spatial and time resolution. These detectors can provide the muon Level-1 trigger system with online track segments of good angular resolution to confirm that muon tracks originate from the IP, considerably reducing the end-cap fake triggers. At the same time, the ATLAS muon system will maintain the full acceptance of its excellent muon tracking at the highest LHC luminosities expected.
Two different detector technologies mounted on NSWs allow to satisfy these requirements: the small-strip Thin Gap Chambers (sTGCs), primarily devoted to the Level-1 trigger function, and Micromegas detectors (MMGs), dedicated to precision tracking. MMGs chambers have exceptional precision tracking capabilities, crucial to maintain the current ATLAS muon momentum resolution in the high background environment of the upgraded LHC. They can, at the same time, confirm the existence of track segments found by the muon end-cap middle station (Big Wheels) online. The sTGC also has the ability to measure offline muon tracks with good precision, so the combination of the two detectors forms a redundant system for triggering and tracking both for online and offline functions.
The impact of the installation of the NSWs on Physics has been studied, mainly on the efficiency of the low single muon trigger (with a threshold of 20-25 GeV) because of its great importance for the future ATLAS physics program. Indeed, Higgs production in pp collisions is mainly due to the gg fusion process in which the p_T of the produced Higgs bosons tends to be low. In the process $H \to WW^* \to l\nu l\nu$, leptons from the W decays are also of low p_T, particularly the subleading lepton, resulting in the loss of a large fraction of the Higgs signals if the Level-1 pT thresholds had to be raised (e.g.to $40\,\mathrm{GeV}$) at high luminosity in order to keep the Level-1 rate manageable. A different possible solution would be not to use the end-cap trigger and to restrict the low p_T single muon trigger to only the barrel region, but the drawback is that the acceptance for the process $H \to WW^*$ would be reduced to about 60%. This is not the only example where the low-p_T muon trigger is important, e.g. $H \to \tau\tau$ has leptons in the final state for which this trigger

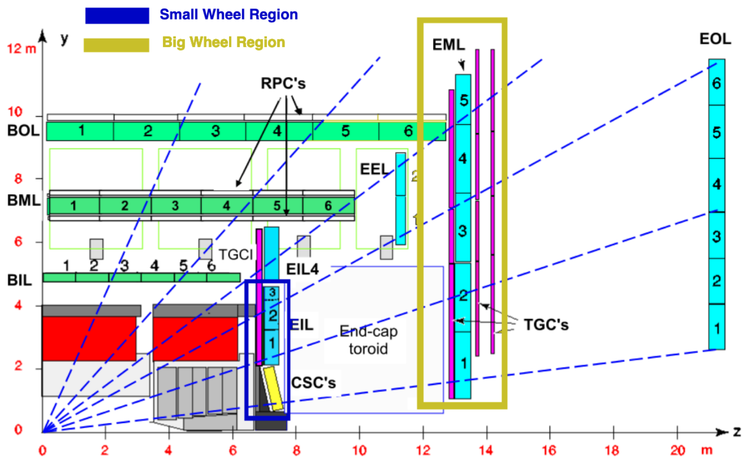


Figure A.1: A z-y view of 1/4 of the ATLAS detector. The blue boxes indicate the end-cap MDT chambers and the yellow box in the Small Wheel area the CSC chambers. The green boxes are barrel MDT chambers. The trigger chambers, RPCs and TGCs, are indicated by the outlined white and the magenta boxes. The detector regions of the Small Wheel and Big Wheel are also outlined in blue and ocra yellow, respectively. [90]

begins crucial. Within the acceptance $|\eta| < 2.5$, the fraction of events with the leading muons having p_T above $25\,\mathrm{GeV}$ is 60% whereas this fraction goes down to 28% for a threshold at $40\,\mathrm{GeV}$. Table A.1 shows a comparison of efficiency from a simulation study for the present detector and the upgrade with NSW for WH with $H \to b\bar{b}$ and for $H \to WW$, along with the efficiencies when restricting the single lepton trigger within the barrel region only.

L1MU threshold (GeV)	$H \to b\bar{b}$ (%)	$H \to WW$ (%)
p_T >20	93	94
p_T >40	61	75
p_T >20 barrel only	43	72
p_T >20 with NSW	90	92

Table A.1: The efficiency for $pp \to WH$ associated production, where $W \to \mu\nu$ and two decay modes $H \to b\bar{b}$ and $H \to WW \to \mu\nu qq'$ are considered. [90]

A.1 Layout of the Wheels

Two NSWs have been built, NSW-A and NSW-C to be positioned in the forward direction on either sides of the interaction point (IP), and equipped with a total of 128 MMG and 128 sTGC chambers. Each wheel of the proposed new detector system consists of 16 detector planes in two multi-layers, comprising four small-strip TGC and four Micromegas planes. The arrangement in each multi-layer consists in a quadruplet of sTGC - MMG -MMG - sTGC, as to maximize the distance between the sTGCs of the two multilayers. As online track hits are reconstructed with limited accuracy, increased distance between detector multilayers leads to an improved online track segment angle reconstruction resolution. The choice of eight planes per detector was dictated by the need to provide a robust, fully functional detector system over its whole lifetime. With eight planes per detector, tracks will be reconstructed reliably and with high precision. In addition, the NSW is expected to operate for the whole life of the ATLAS experiment; the large number of planes will ensure an appropriate detector performance even if some planes fail to work properly.
The two NSW detector technologies also complement each other for their corresponding primary functions. sTGC may contribute to offline precision tracking, as they are able to measure track hits with a resolution better than $150\,\mu\mathrm{m}$. For triggering, experience has shown that redundancy is highly important in the forward direction at high luminosities; the MMG detectors will be employed as a trigger in addition to the sTGC to provide improved redundancy, robustness and coverage of the forward trigger.
Going back to each wheel sectors, they are 16, 8 small and 8 large. Each sector is composed by 2 identical MMG wedges sorrounded by two sTGC wedges in such a way that a muon coming from the LHC interaction region encounters the following detector sequence: 4 sTGC layers, 8 MMG layers and finally 4 more sTGC layers. A scheme of the general structure of a New Small Wheel is shown in Figure A.2. The chambers are organised in wedges: each small wedge consists of an SM1 and an SM2 chamber, each large wedge of an LM1 and an LM2 chamber (Figure A.2 right). Two identical wedges, mounted on opposite faces, form a MMG double-wedge. A NSW sector is finally obtained with the addition of two sTGC wedges, one per side of the double-wedge.
The sTGCs wedges consists of three modules in the *r* direction with four planes each. The modules in the two sTGC wedges of each sector are arranged so that no dead region exists in projective geometry with respect to the IP. Figure A.3 shows the sTGC and MMG detectors inside the mechanical support structure of the NSW.

In the following subsections the details of detector technology is described for both

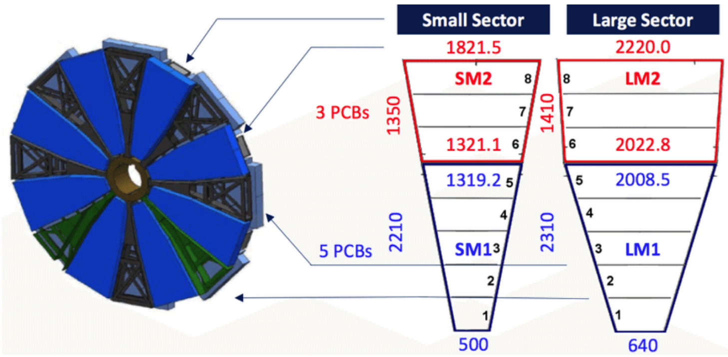
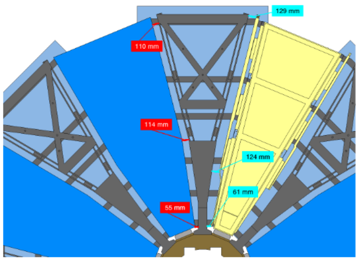
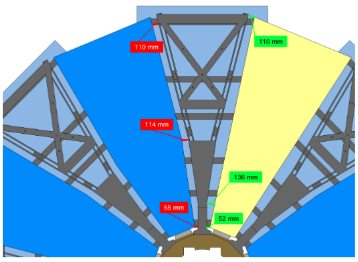


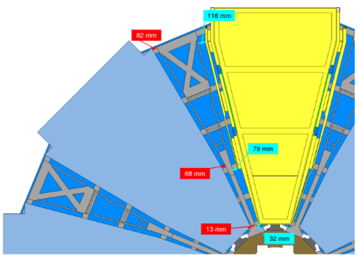
Figure A.2: (Left) General structure of NSW. (right) Geometry of the MMG wedges for small and large sectors, and dimensions (quotes in mm) of the two Micromegas chambers in each wedge. [18]



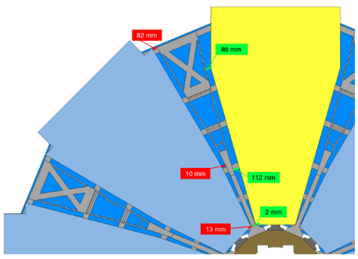
(a) sTGC (small sector).



(b) MM (small sector).



(c) sTGC (large sector).



(d) MM (large sector).

Figure A.3: Layout of the sTGC and MMG detectors with services (yellow) and the allowed envelopes (blue). The distances at various points to the NSW structure are indicated for both. [90]

MMGs and sTGCs separately.

A.1.1 MicroMegas detector

The Micromegas detectors, proposed in the 90s [79], belong to micro pattern gaseous detectors generation, characterized by optimal space resolution together with a good multi-track separation capability in the high rate environment. This kind of detectors profited from the development of photolithographic techniques for the design of high-granularity readout patterns and, in parallel, from the development of specialised front-end electronics able to readout an increased number of channels [1, 97].
A schematic view of the configuration and of the MMG operating principle is shown in Figure A.4. The main components of the detector (from the top to the bottom of the figure) are:

- a planar metallic cathode kept at a negative voltage;
- a $5\,\mathrm{mm}$ thick gap (drift gap) with a relatively low electric field (about $0.6\,\mathrm{kV/cm}$) where a ionising particle produces electron-ion pairs in the gas and electrons drift towards the mesh;
- a thin metallic micro-mesh kept at ground potential;
- a $120\,\mu\mathrm{m}$ thick gap (amplification gap), whose thickness is maintained by insulating pillars standing on the anode plane, with a high electric field $(40 - 50\,\mathrm{kV/cm})$ where electrons passing through the mesh give rise to avalanches;
- a segmented anode with resistive strips (about $400\,\mu\mathrm{m}$ pitch) kept at a positive voltage and supersimposed to the readout strips, where the signals are capacitively induced. The precision coordinate is orthogonal to the strip direction.

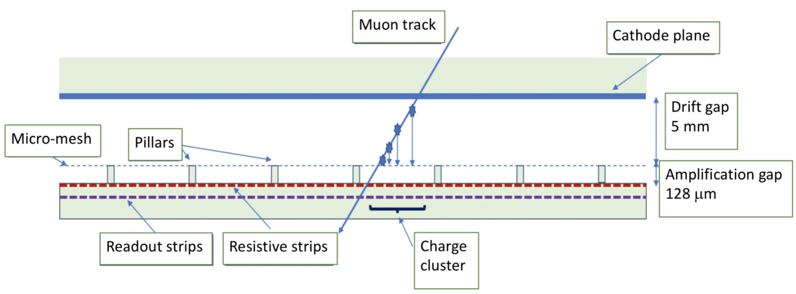


Figure A.4: Schematic view of a Micromegas layer. A muon track and its ionisation are also shown. [18]

The Micromegas can be operated with different gas mixtures. A gas mixture of 93% Ar and 7% CO2 is the standard mixture for ATLAS and has been used for the chamber validation, even though other gas mixtures based on Ar/CO2 in different proportions or with the addition of a small percentage of isobutane, have been tested. 93/5/2% Ar/CO2/isobutane has been adopted for ATLAS Run-3, since it shows more stable current conditions for almost all the sections, the isobutane being a gas quencher that shuts off discarges due to its roto-vibrational degrees of freedom absorbing photons energies.
The high ratio between the electric fields in the amplification and drift regions has two important consequences: the mesh turns out to be almost transparent to the drifting electrons; most of the ions produced by the avalanches in the amplification region reach the mesh and are evacuated in a relatively short time, of the order of $100\,\mathrm{ns}$, reducing the distortion of the electric field in the drift region. This is particularly important for applications in high-rate environments.
Several test-beams done on small-size prototypes showed that a Micromegas detector is able to reconstruct the position to reach a spatial resolution of $100\,\mu\mathrm{m}$ for all incidence angles [22, 23]. Moreover it was shown that a very good time resolution could be reached, at the level of $10\,\mathrm{ns}$.

The Italian collaborations (INFN sites of LNF, Cosenza, Roma1, Roma3, Pavia, Lecce and Napoli) have directly participated in the construction of SM1 modules and their layout is explained in more details below.
The active area of SM1 chambers is a trapezoid with $221\,\mathrm{cm}$ height, a larger base of $132\,\mathrm{cm}$ and a smaller base of $50\,\mathrm{cm}$. The total thickness of the structure is slightly less than $8\,\mathrm{cm}$. Each panel surface is segmented in five trapezoidal printed circuit boards (PCBs) of different sizes with a maximum height of $45\,\mathrm{cm}$.

SM1, as the other modules, are structured as quadruplets consisting of five panels with four sensitive layers, as shown in Figure A.5. Each panel is about $11\,\mathrm{mm}$ thick, and they are separated by $5\,\mathrm{mm}$ thick gaps. The two external panels and the one in the centre provide the four cathode planes (one plane in each external panel and one plane in each face of the central one) and they are called *drift panels*. In order to create one drift gap in the external panels and two drift gaps in the central one, micro-mesh foils are stretched on them. The other two panels, the second and the fourth, are the readout panels (*RO* panels from now on). On both faces of each of them a total of five readout PCBs are glued, providing the anode planes, including the insulating pillars. The two RO panels, named *eta* and *stereo*, are different. In the eta panels the resistive and the readout strips run parallel to the trapezoid bases; therefore the precision coordinate is perpendicular to the trapezoid bases, and corresponds to the η coordinate in the experiment coordinate system. In the stereo panels the resistive and readout strips run with small angles ($\pm 1.5°$ in the two faces of the panel) with respect to the direction of the bases. This angle allows reconstruction of ϕ, the azimuthal coordinate in the experiment coordinate system. The mesh is integrated on the drift panels, glued on $5\,\mathrm{mm}$ thick frames around the cathode. The amplification gap is created when a drift panel is coupled to a RO panel. The mesh touches the pillars located on the RO panels, and is therefore kept at the desired distance from the anode. The main advantage of this setting, called floating mesh, is to have a structure that can be re-opened in case of need, for example to eliminate defects on the anode plane or to improve cleanliness.

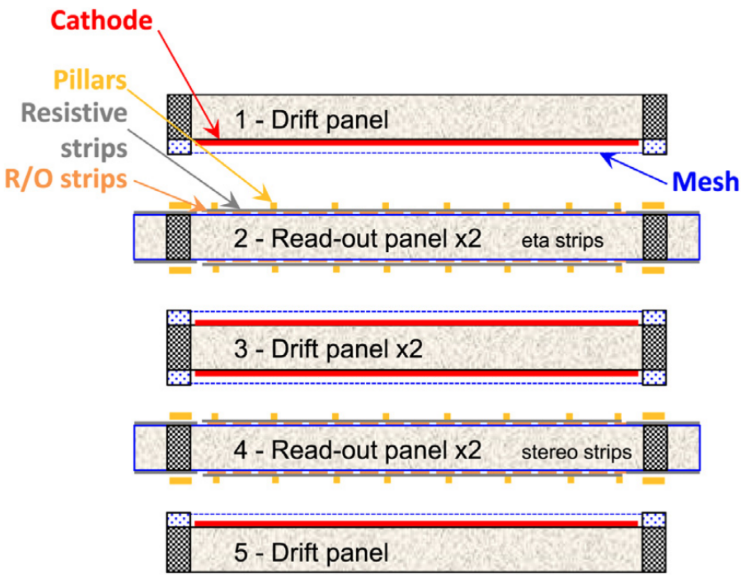


Figure A.5: A schematic layout of the MMG five panels in a quadruplet. [18]

A.1.2 small-strip TGC detector

sTGC basic chamber structure consists of a grid of $50\,\mu\mathrm{m}$ gold-plated tungsten wires with a $1.8\,\mathrm{mm}$ pitch sandwiched between two cathode planes at a distance of $1.4\,\mathrm{mm}$ from the wire plane. The cathode planes are made of a graphite-epoxy mixture sprayed on a $100\,\mu\mathrm{m}$ thick garolite plane, behind which there are on one side strips (that run perpendicular to the wires) and on the other pads (covering large rectangular surfaces), on a $1.6\,\mathrm{mm}$ thick PCB with the shielding ground on the opposite side. sTGC's layout is shown in Figure A.6. The strips have a $3.2\,\mathrm{mm}$ pitch, much smaller than the strip pitch of the ATLAS TGC, hence the name *small*

TGC for this technology.
The layouts of the chamber designed for the NSW consists in two quadruplets $35\,\mathrm{cm}$ apart in z. Each quadruplet contains four TGC's, each TGC with pad, wire and strip readout. The pads are used to produce a 3-out-of-4 coincidence to identify muon tracks roughly pointing to the interaction point. They are also used to define which strips are to be readout to obtain a precise measurement in the bending coordinate, for the online muon candidate selection. The azimuthal coordinate, where only about $10\,\mathrm{mm}$ precision is needed, is obtained from grouping wires together. The charge of all strips, pads and wires are readout for offline track reconstruction.

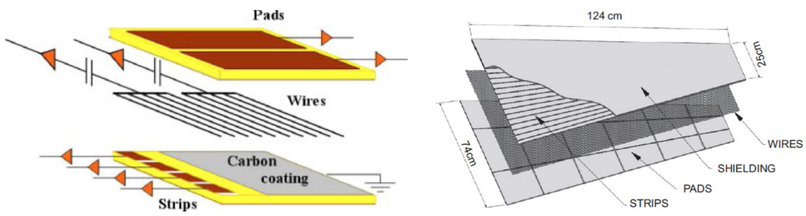


Figure A.6: A schematic layout of the sTGC's internal structure. [90]

The high background of the HL-LHC environment necessitated modifying the TGC technology in order to achieve a very good position resolution at high count rates (approximately $100\,\mu\mathrm{m}$). The main emphasis during the development stage was on achieving this performance with the smallest possible number of electronic channels. It is known that space-charge effects are not a limiting rate element for TGC. However the rate capability is limited by the use of the resistive coating in the cathodes, which reduces, under high irradiation, the effective operating voltage of the detector in areas far from the ground contacts. As a consequence a low surface resistivity coating has been used, and the capacitance between the strips/pads and the cathode has been increased to keep the same transparency for fast signals. This optimization, combined with a reduction of the HV and ground decoupling resistors is sufficient for maintaining high efficiency for minimum ionizing particles in large surface detectors subject to detected rates of up to $20\,\mathrm{kHz/cm^2}$ over the full surface.

A.2 NSW TDAQ system

The NSW trigger will enable to drastically reduce the trigger rate in the end-cap by reducing fake triggers coming from non-pointing tracks. The trigger is based on track segments produced online by the sTGC and MMG detector. The NSW trigger system provides candidate muon track segments to the new TGC Sector Logic which uses them to corroborate trigger candidates from the Big Wheel TGC chambers. In Figure A.7 track *B* is rejected because the NSW does not find a track coming from the interaction that matches the Big Wheel candidate, as well as track *C*, rejected because the NSW track does not point to the interaction point; track *A* is the only one having a correct correspondence in both wheels and is therefore taken. The Sector Logic then sends Level-1 trigger candidates to the ATLAS Muon Central Trigger logic.

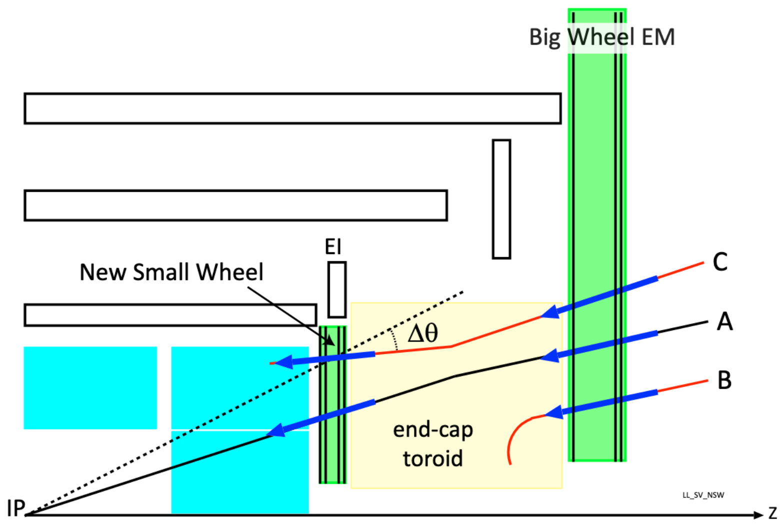


Figure A.7: Example of the Muon End-cap trigger: only track *A* is accepted because it is confirmed by both the Big Wheel and the New Small Wheel. [90]

Wherever possible the two technologies, sTGC and MMG, use the same building blocks to construct their trigger and data acquisition (TDAQ) systems.
The background radiation levels at the NSW, although not as high as those encountered in the Inner Tracker, are sufficiently high that they affect the design of the on-detector electronics as well as the choice and radiation certification of custom and off-the-shelf components. Extrapolating to the expected luminosity of 5×10^{34} and for 10 years of operation, the Total Ionizing Dose (TID) at the highest pseudo-rapidity region of the NSW will be of the order of 0.5 mRad (with some uncertainty factored in). Modern deep sub-micron CMOS processes are generally

sufficiently radiation tolerant at such levels.
In the fall of 2009 an effort was launched to develop a custom front-end Application Specific Integrated Circuit (ASIC) that could be used to read both the sTGC and the MM detectors of the ATLAS new Small Wheels (NSW). The ASIC is named *VMM* [87]. Both NSW detectors require precision amplitude measurement for position determination by charge interpolation. In order to mitigate the deterioration of the spatial resolution for inclined tracks in the MMG technology these detectors will be operated as a so-called *TPC* where the track is reconstructed by measuring the drift time of charges arriving on individual strips as is done in large Time Projection Chambers (TPC). Therefore, time measurement with precision of 2 ns, along with the amplitude measurement, are needed and the ASIC allows to perform both. It provides the peak amplitude and time with respect to the bunch crossing clock, (or other trigger signal) along with triggers, direct timing outputs, sparse and derandomized readout, and multiplexing. The limited bandwidth of the readout link requires on-chip zero suppression. For reliable operation in the ATLAS environment, radiation tolerance is also required as anticipated above.
Each VMMs is composed of 64 front-end channels each providing a low-noise charge amplifier (CA), a shaper with baseline stabilizer, a discriminator with trimmer, a peak detector, a time detector, some logic, and a dedicated digital output for ToT (Time-over-Threshold) or TtP (Time-to-Peak) measurements. Channels share with each others the bias circuits, a temperature sensor, a test pulse generator, two 10-bit DACs for adjusting the threshold and test pulse amplitudes, a mixed-signal multiplexer, the control logic, and the ART (Address in Real Time) which consists of dedicated digital outputs (flag and address) for the first above-threshold event. A schematic view of the VMM is shown in Figure A.8.

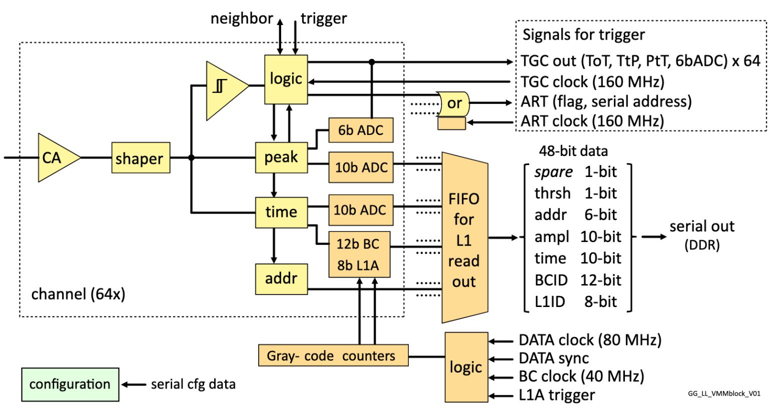


Figure A.8: Block diagram of the VMM ASIC. [90]

For the MMG, a design for cooling each doublet of planes is considered. The electronics heat is carried away by a cooling channel running down between the front-end cards of the doublet on one side of the sector, and out between the cooling cards on the other side. Working with a temperature change of the cooling liquid of $3\,\mathrm{C}$ the required waterflow in one channel is $21.5\,\mathrm{ml/s}$ or $77.4\,\mathrm{ml/s}$. The heat will be transported from the chip surface to the cooling channel by a thermally conductive ceramic foam pad. Starting with a cooling water temperature of $17\,\mathrm{C}$ which gets warmed up to $20\,\mathrm{C}$, this design will keep the temperature of the front-end chips safely below $30\,\mathrm{C}$, the maximum tolerable temperature.

The two detector technologies will use the same trigger processor cards in the ATCA standard (Advanced Telecommunications Computing Architecture) as the platform for their trigger logic in USA15 (the cavern close to the ATLAS detector at Point 1, dedicated to host all the electronics), but with different firmware and numbers of input links.
At each bunch crossing, the sTGC trigger electronics finds local tracks that point, with $1\,\mathrm{m}$ precision, to the Big Wheel to corroborate its coincidences. In order to place the track finding and extrapolation logic off-detector, a 3-out-of-4 coincidence of pads in each of the 4-layer quadruplets is used to choose the relevant bands of strips to be sent off-detector. This substantially reduces both the required bandwidth and the amount of centroid and track finding logic. To reduce the number of pad channels, the layers are staggered by half a pad in both directions to make logical towers one quarter the area of a pad. The width of the pads in ϕ decreases with lower radius to handle the increased background rate.
As a baseline, the MMG and sTGC will independently compute track vectors. Both sTGC and MMG trigger modules belonging to a 1/16th sector will be in the same ATCA trigger crate which has a fully interconnected backplane, i.e. all modules are connected to every other module by several high-speed, low latency differential pairs. In this way MMG vectors will be sent to a final merging stage of the sTGC trigger and vice versa. The output of the merging stage is sent to the Sector Logic. Since both merging stages would output the same merged list of vectors, only one must actually be connected to the Sector Logic. This provides redundancy in the case of failure of one of the modules. Each vector sent to the merging stage is tagged with a 'quality' flag that is used by the merging logic to select the best vectors.
The off-detector readout chain of the ATLAS detectors presently consists of sub-detector specific custom ReadOut Drivers (RODs), usually FPGA-based, which output event fragments to the ReadOut System (ROS) via the ReadOut Links (ROLs). The ROLs are point-to-point custom optical links that cross the boundary between the sub-detector domains and the TDAQ domain. The Readout Architecture Upgrade Working Group, with representatives from all detectors and TDAQ, has considered a new architecture that better promises to take advantage of future

advances in electronics and information technology. The new architecture reduces the amount of custom electronics and links in favour of industry standard equipment called FELIX (Front End LInk eXchange). Two important advantages of the architecture based on FELIX are: 1) minimization of the amount of special-purpose detector-specific hardware and associated firmware and therefore the flexibility of general purpose, industry standard equipment, and 2) the ability to easily scale in power and be upgraded with new technologies to meet future needs.

A.3 NSW SM1 commissioning

In this section commissioning activities regarding chambers assembly and quality controls (QA/QC) are described for SM1, carried out at LNF in Frascati (RM).
I directly participated in these activities for my Qualification Task, a service activity that the ATLAS collaboration requires for being included in its authors list.

A.3.1 Panels preparation

An initial cleaning procedure for both Drift and RO panels was designed, in order to minimize the problem of unstable conditions of the chambers when HV was applied due to the remnants from the industrial production of PCBs. Before washing, a fine-grade sand paper was passed over the wet mesh of the Drift panel in order to smooth the mesh surface and remove possible imperfections which might induce sparks in the amplification gap. The cleaning was performed by washing the panel when placed in vertical position in a custom-made large basin A.9a. Subsequently, a clean room tissues and a flow of nitrogen were used to remove the water drops from assembly holes, rims and gas tubes. Then each panel was moved to the drying station, a custom-made structure A.9b able to host the five panels needed for a MM quadruplet in a vertical position. The drying station was equipped with a ventilation system to filter the air and keep the temperature in the range $38° - 40\,°C$. The panels were left to dry for a minimum of two days before chamber assembly.
Then the four panels of a quadruplet are assembled together in a clean room and set in a horizontal position on a granite table to start the QA/QC.
It consisted of the following tests:

1. measurements of planarity and thickness;
2. measurement of gas tightness;
3. measurement of the PCB alignment among the four RO layers;
4. assessment of the high-voltage stability.

(a) (b)

Figure A.9: Washing basin with a hanging drift panel (a). Drying station with RO panels inside it (b). [18]

Once the module has successfully passed these checks, it undergoes a cosmic ray test, otherwise the chamber is first reopened and fixed.

Planarity and thickness Planarity and thickness measurements have been taken by positioning the chamber horizontally on specific supports on a granite table, inside the clean room. A laser tracker pointed to a retro-reflective target that was moved across the chamber surface by the operator. A height map was obtained for each side of the chamber and a planar fit was performed. Figure A.11 shows the measured points and the interpolated surface for both sides of a chamber. The planarity of the chamber is extracted from the RMS of all the measurements. The thickness is evaluated by calculating the difference between the two maps from both sides of the same chamber and calculating the mean value of all the measurements. Figure **??** shows the thickness and the planarity values for all the SM1 chambers.

Gas tightness The overall gas leakage test of each chamber, after its assembly, was performed in the clean room always on the granite table positioning the chamber horizontally. The chamber was over-pressurised in a static way, by connecting the gas input line of the chamber to a syringe and the gas output to a pressure sensor. The value of the initial pressure inside the chamber was taken as reference;

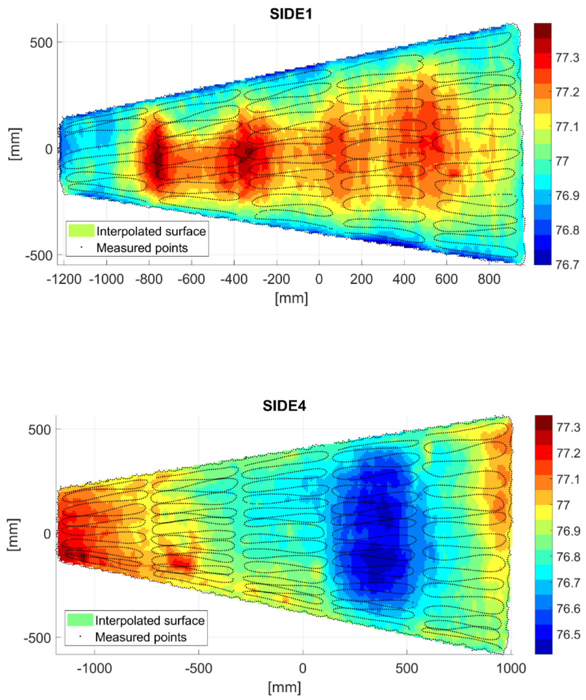
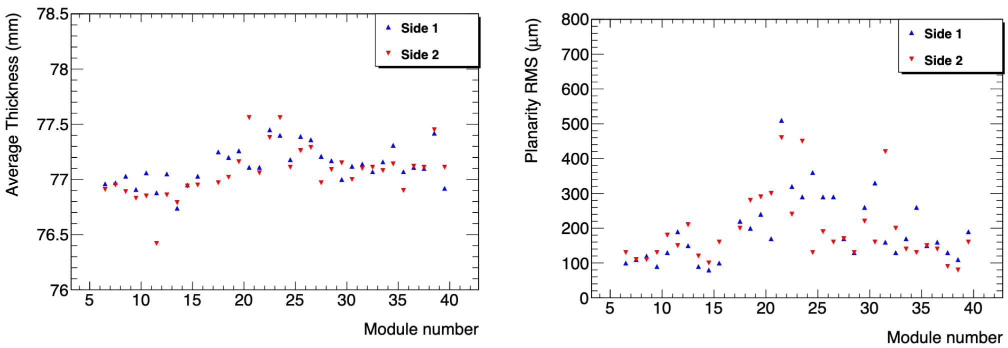


Figure A.10: Results of the planarity measurements of the two faces of a chamber. [18]



(a)

(b)

Figure A.11: Thickness measurements (expressed in mm) for all the SM1 chambers (a). Planarity measurements (expressed in μm) for all the SM1 chambers (b). [18]

then air was injected into the chamber with the syringe, until an over pressure of $3\,\mathrm{mbar}$ was reached. The pressure variation ΔP was then monitored to estimate the pressure drop over time. To extract a measurement of the leakage in mbar/hour, the

ΔP over time data during the pressure drop period was fitted with a linear function, as a good approximation of the exponential behaviour within the 15–20 minutes of the measurement duration. The gas leak is then calculated as the angular coefficient of the fit straight line and it is shown in the summary graph in Figure A.12. The clean room temperature was monitored during this measurement.
All the SM1 chambers had leakage rates well below 5 times the limit set, that was -0.64 mbar/hour.

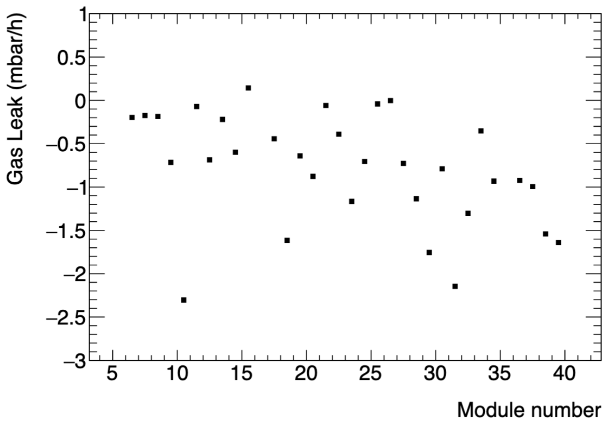
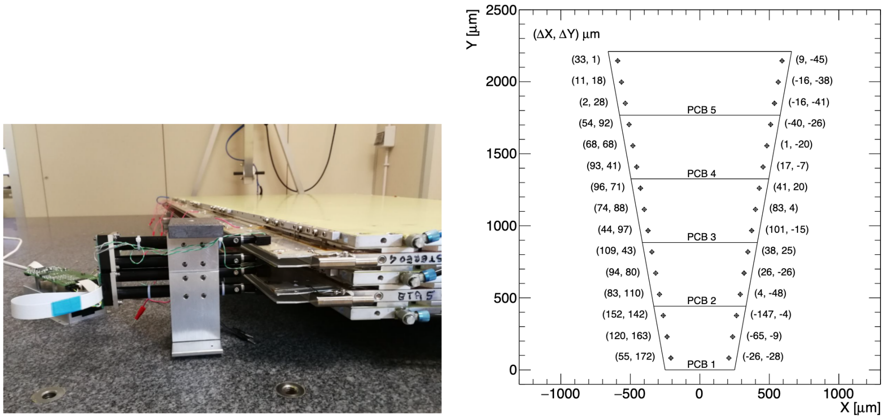


Figure A.12: Gas leakage measurements (expressed in mbar/hour) for all the SM1 chambers. [18]

PCBs alignment The measurement of the panel-to-panel PCB alignment was performed in CR-1 using a 4-rasfork tool (a modification of the rasfork technique, a micropattern coded mask read by an optical system, developed in Saclay) that is able to measure the misalignment of the two corresponding PCBs in the two read-out panels of a quadruplet, by exploiting the localisation precision of a *Rasnik masks*. The tool configuration is shown in Figure A.13a. A map of the measured misalignment (ΔX, ΔY) is shown in Figure A.13b for one of the tested chamber (the Y coordinate represents the precision coordinate η in the ATLAS coordinate system, while X is the orthogonal coordinate in the transverse plane 2.2.1).
HV stability Preliminary tests for HV stability were conducted in air, during the assembly, and with Ar/CO2, before the closure of the chambers inside the clean-room; then a final test was performed on the cosmic ray stand. Each chamber was connected to 40 different HV cables: namely each of the five RO PCB in the four gaps, being divided into two HV parts, one per side. With the setup used it was possible to power each of the 40 HV sections of the chamber independently. The four cathode channels, one per drift layer, were also supplied with the nominal -300 V value.
The HV test was performed flushing the chamber with a 20 l/h gas flow of a $93 : 7\%$



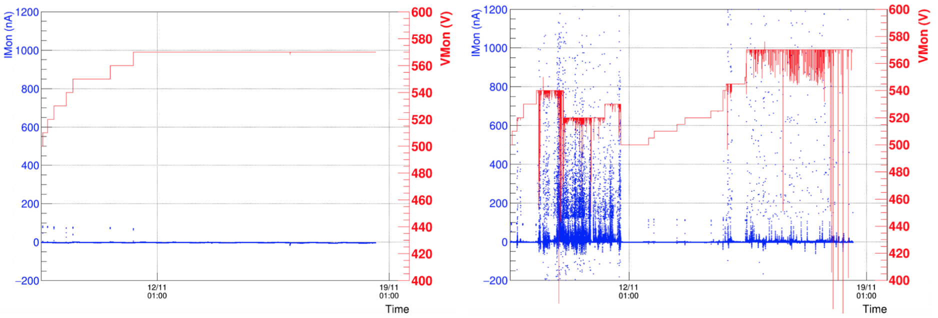
(a) (b)

Figure A.13: Rasfork measurement on the granite table (a). A measurement showing the position of the Rasnik masks in the X and Y coordinates (b) for a chamber. [18]

Ar/CO2 gas mixture. Increased gas flow reduces the relative humidity inside the gas gap, minimising the contamination of the gas mixture with water. The HV ramp-up was started when the gas relative humidity level reached a value 10%. Then each section was ramped up with steps of 5-10 V, from a value of 400 V to 570 V.

The values of the more interesting parameters, the monitored HV (VMon) and the monitored current (IMon), were constantly recorded by a DCS code developed at LNF.

For some HV sections it was not possible to reach the nominal value of HV_{max} = 570 V and the limits for them were set based on the current drawn and on the spark rate, defined as the frequency at which IMon goes above the defined current threshold of 50 nA. Being the current sampled every second, the spark rate is defined as the number of seconds with a current exceeding the threshold in one minute. When the spark rate exceeded 6/minute, the HV was lowered until a stable condition was reached. Figure A.14 shows an example of a good (left) and a bad (right) section, where sparks already begin at HV = 500 V and the chamber cannot reach a stable condition. To meet the ATLAS HV acceptance requirement for an SM1 chamber, at least 85% of the sections were required to be stable at the nominal HV value of 570 V.



(a) (b)

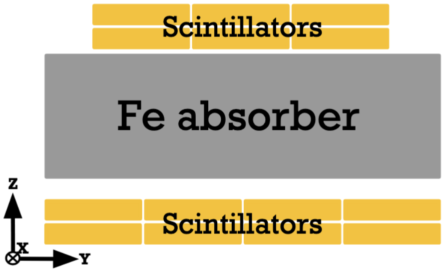
Figure A.14: VMon (blue) and IMon (red) versus time for two sections: (a) a good section where the HV ramp up to HV$_{max}$ = 570 V is accompanied by no sparks at all and (b) a bad section with so many sparks that the nominal voltage could not be reached. [18]

A.3.2 Cosmic-ray test stand

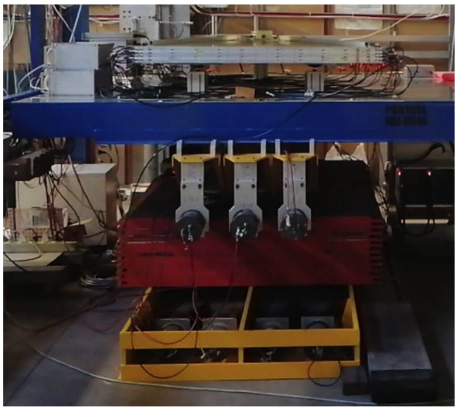
The final QA/QC step is the test of detector efficiency at the cosmic-ray stand (CRS). The CRS is a multi-layer structure where the trigger is provided by four planes of scintillators, extending over an area of about $1\,\mathrm{m}^2$, covering the full chamber length in the η coordinate.
Starting from the top to the bottom (Figure A.15b), it consists of:

- the base where the SM1 chamber is collocated;
- two upper planes of plastic scintillators for the trigger system, where each plane is formed by 3 $20 \times 150\ cm^2$;
- iron absorber, formed by a 35 cm thick stack of iron plates, providing an energy cut of approximately 0.6 GeV;
- two lower planes of plastic scintillators as above, but in this case there are four vertical pairs of slabs.

Scintillators are coupled to photomultipliers (PMT) by light guides. The trigger signal is issued if there is a coincidence between a vertical pair of slabs above and a vertical pair below the iron absorber (Figure A.15a). The overall trigger rate is around 50 Hz.
The relative humidity of the gas in the chamber was monitored by a humidity sensor and the relative humidity was maintained below 10% during the whole test. For the chamber performances meausurements the signal is read in the following way: adjacent hit channels are grouped in clusters and the cluster charge (Q_{CLU}) is



(a)



(b)

Figure A.15: The trigger system scheme (a). A frontal picture of the cosmic-ray stand. Trom top to bottom: the SM1 chamber, the upper scintillator slabs, the iron absorber and the lower scintillator slabs. [18]

given by the sum of the charge collected by each channel composing the cluster. The cluster centroid is given by the charge-weighted mean of the positions of all channels in the cluster.

In order to compute the efficiency in a given layer, a track is reconstructed in the other three layers, used as control layers. Only events with one and only one cluster in each control layer are retained. Reconstructed clusters in the layer under study are compared to the position of the extrapolated track in that layer. Defining N_i^{trk} as the number of tracks extrapolated in the i-th bin and N_i^{CLU} as the sub-sample for which a cluster is found, the efficiency of the i-th bin in a layer is defined as:

$$\epsilon_i^{layer} = \frac{N_i^{CLU}}{N_i^{trk}} \tag{A.1}$$

By using the information from the stereo layers, it is possible to obtain a 2D efficiency map. If the layer under examination is an eta layer (layers 1 or 2), the ϕ coordinate of the extrapolated track, corresponding to the y-axis in the plots, is given by:

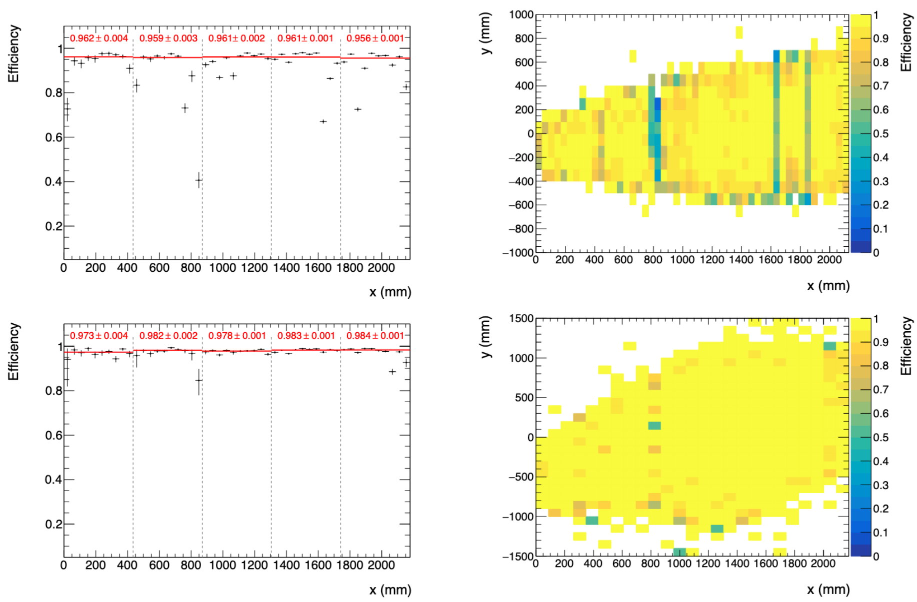
$$\phi = \frac{l_3 - l_4}{2 \sin\theta} \tag{A.2}$$

where l_3 and l_4 are the positions of the centroids in the stereo layers, and θ is the angle of inclination of the stereo strips ($1.5°$).

If the layer under examination is a stereo layer (layers 3 or 4), the ϕ coordinate is given by:

$$\phi = \frac{l_2 \cos\theta - l_{stereo}}{\sin\theta} \tag{A.3}$$

where l_2 is the position of the centroid in the second eta layer while l_{stereo} is the position of the centroid in the other stereo layer. In this case a zero-distance approximation in the vertical coordinate (Z) between l_2 and l_{stereo} is used.
An example of 1D efficiency as a function of the precision coordinate (η) for two different layers, one eta and one stereo, is shown in Figure A.16a. This is a good case, where the efficiency is around 97%. The few low-efficiency spots are related to problems in the front-end cards due to the RO connectors deterioration. The corresponding 2D efficiencies are also shown in Figure A.16b.



(a) (b)

Figure A.16: (a) 1D efficiency at HV$_{max}$ = 570 V as a function of the position for: (top) eta layer 1 with all the sections operating, (bottom) stereo layer 3 of the same SM1 chamber with all the sections operating. (b) 2D efficiency for the same layers. [18]

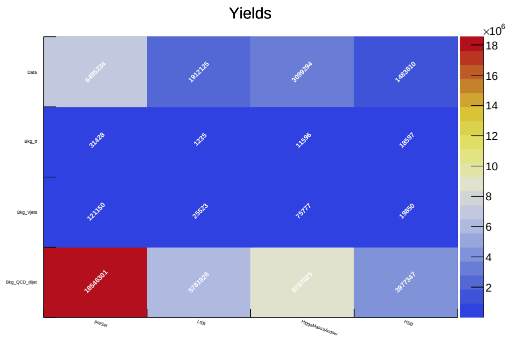
For all the detectors built, the overall efficiency is above 90% when operating at HV$_{max}$ = 570 V using a 93/7 % Ar/CO2 gas mixture. Moreover, the study of chambers operating with a gas mixture including a 2% of isobutane has shown a clear improvement in the high-voltage stability, that's why the latter have been adopted as the nominal choice for the MMG for Run-3.

APPENDIX B

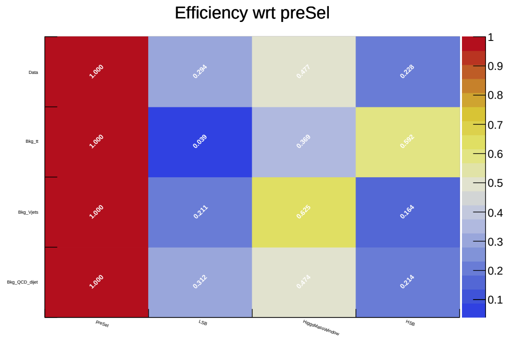
YXH analysis insights

B.1 Background composition from MC

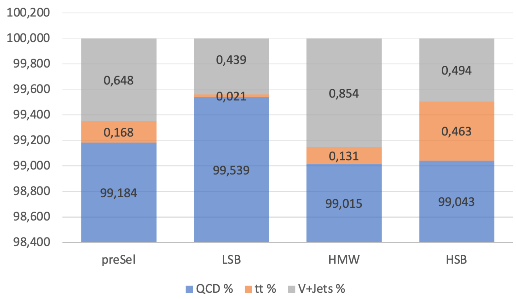
Most of the background in the fully hadronic final state consists in QCD multi-jet processes. Other contributions come from V(Z/W)+jets and $t\bar{t}$ productions, but their percentage is very low.
In this appendix, from Figure B.1 to B.4, yields and efficiencies at $139\ fb^{-1}$ after the selections described in **??** are calculated for the three MC background samples along with data.
Even though this analysis is blinded it is safe to look at the total data events in SR, since the eventual excess should not be visible on a simple counting-based experiment.



(a)

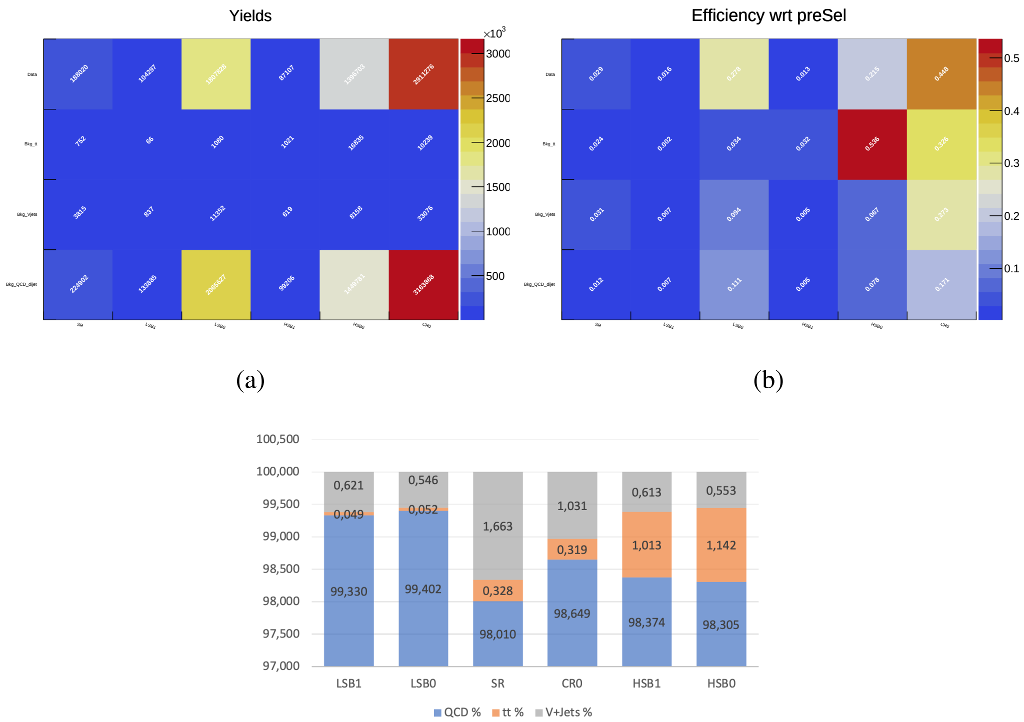


(b)



(c)

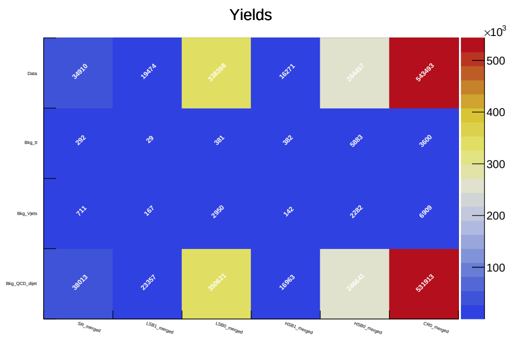
Figure B.1: Yields (a) and Efficiencies with respect to PreSelection (b) for Data and Monte Carlo background samples, after dividing higgs candidate mass windows. The background relative composition is shown in (c).



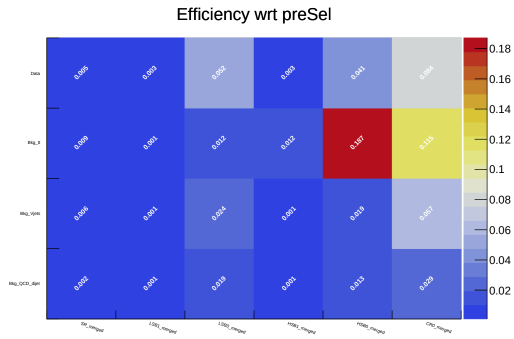
(a) (b)

(c)

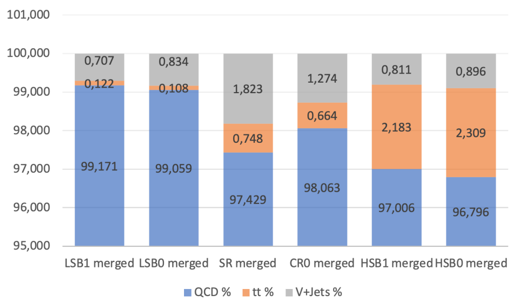
Figure B.2: Yields (a) and Efficiencies with respect to PreSelection (b) for Data and Monte Carlo background samples, after applying $D_{H_{bb}}$cut. The background relative composition is shown in (c).



(a)

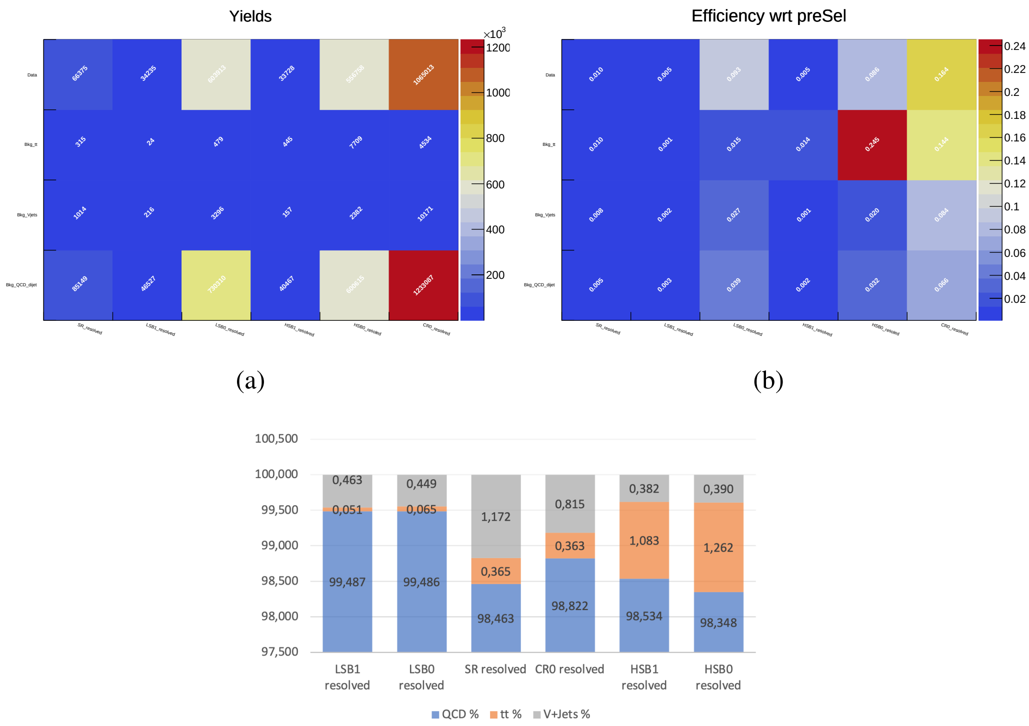


(b)



(c)

Figure B.3: Yields (a) and Efficiencies with respect to PreSelection (b) for Data and Monte Carlo background samples, after applying the Two-prong merged selection. The background relative composition is shown in (c).



(a) (b)

(c)

Figure B.4: Yields (a) and Efficiencies with respect to PreSelection (b) for Data and Monte Carlo background samples, after applying the Two-prong resolved selection. The background relative composition is shown in (c).

B.2 Single Higgs contributions in SR

In the background composition described above, contributions from single Higgs production have not been considered, but in principle they may be present in SR. In this subsection the expected yields for the SM Higgs processes are compared with the other expected background already discussed and they are found to be negligible.

For each background process the selection efficiencies ϵ_{HSB} and ϵ_{SR} has been computed for the HSB1 and SR. They are defined as:

$$\epsilon_{HSB} = \frac{yields\ in\ HSB1}{yields\ selecting\ m_H > 145\ GeV} \tag{B.1}$$

$$\epsilon_{SR} = \frac{yields\ in\ SR}{yields\ selecting\ 75\ GeV < m_H < 145\ GeV}. \tag{B.2}$$

Moreover, the ratio between the yields of a particular background process and the expected yields of the QCD multi-jet has been computed both in HSB1 and SR.

A summary of the results can be seen in Table B.1 where two tables containing the results of the background studies for HBS (above) and SR (below), both in Two-prong merged and resolved categories are shown. The column "*Double Ratio*" indicates the ratio between the HSB1 background yields/QCD yields and the SR background yields/QCD yields. From these, it is evident that the non-QCD SM background represents a small contribution to the total expected background yields. Among these, in principle, Z+jets and SM Higgs processes give the most significant impact on the total background shape from HSB to SR. Indeed, these two components increase their relative contribution by a factor of ~ 2 in the resolved SR, and a factor of ~ 3 in the merged SR.

B.3 Further studies of the discovery region unblinded results

Figure 4.95 shows a deviation between data and background corresponding to a p-value of 8.8$\times$10^{-3} in the m_X bin of (75.5, 95.5) GeV. The BUMPHUNTER interval in m_Y is determined to be (3424.0, 3805.0) GeV, with the excess being centered at roughly 3.7 TeV.

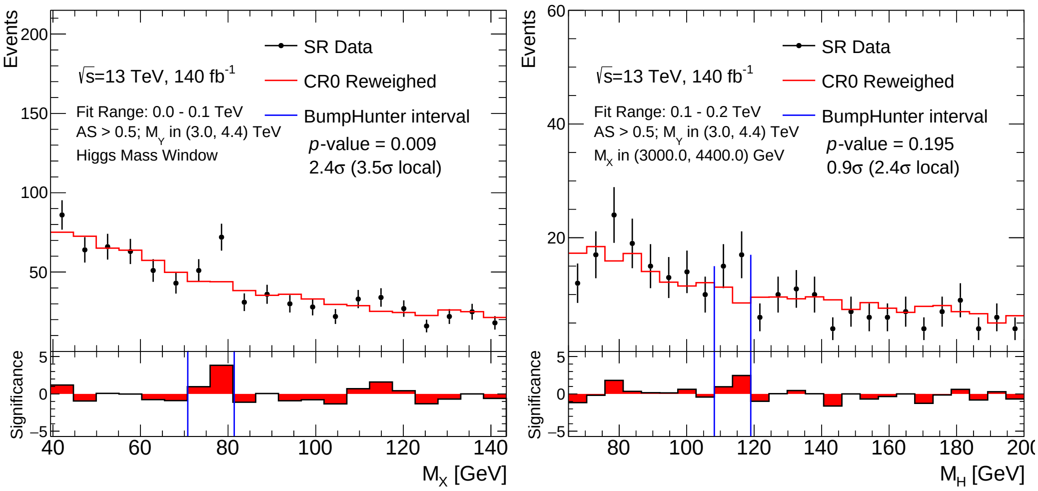
	Background	Efficiency (%)	yields HSB1 bkg/QCD(%)
HSB RESOLVED	W + jets	0.37	0.14
	Z + jets	3.44	0.70
	$t\bar{t}$	2.37	0.59
	QCD	1.05	100
	SM Higgs	12.29	0.06
HSB MERGED	W + jets	0.29	0.28
	Z + jets	1.06	0.53
	$t\bar{t}$	1.97	1.26
	QCD	0.42	100
	SM Higgs	5.32	0.07

	Background	Efficiency (%)	yields SR bkg/QCD (%)	Double Ratio
SR RESOLVED	W + jets	0.31	0.16	1.13
	Z + jets	5.21	1.29	1.85
	$t\bar{t}$	2.81	0.35	0.60
	QCD	1.08	100	
	SM Higgs	17.51	0.11	1.70
SR MERGED	W + jets	0.24	0.33	1.16
	Z + jets	2.45	1.54	2.89
	$t\bar{t}$	2.52	0.77	0.61
	QCD	0.43	100	
	SM Higgs	12.51	0.21	2.97

Table B.1: Tables containing the results of the background studies for HBS (above) and SR (below), both in merged and resolved regimes. The *Double ratio* column indicates the ratio between the yields ratio HSB1 background/QCD and the yields ratio SR background/QCD.

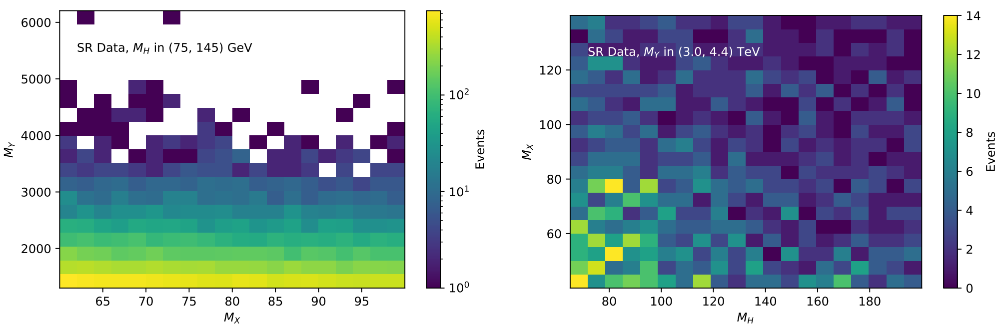
B.3.1 Corresponding m_X and m_H distributions

The possibility of this excess corresponding to a resonant new particle has been investigated, by first selecting an m_Y window of (3.0, 4.4) TeV surrounding it. Figure B.5 shows that in doing so, we observe a corresponding excess in $m_X$of 2.5σ (3.5σ local) significance at ~80 GeV and high bins in $m_H$at ~80 and ~115 GeV at 0.9σ (2.4σ local) significance.
Figure B.3.1 shows the corresponding 2D distributions of $m_Y \times m_X$ and $m_h \times m_X$ in windows surrounding the excess, showing hot spots in these distributions corresponding to the excesses observed in each of the mass variables.



(a) (b)

Figure B.5: BUMPHUNTER fits of m_X (a) and m_H (b), restricting m_Y to a window of (3.0, 4.4) TeV enclosing the original excess. A corresponding excess of 2.5σ (3.5σ local) significance at $\sim$80GeV in $m_X$and high bins at $\sim$80 and $\sim$115 GeV of 0.9σ (2.4σ local) significance in $m_H$are observed.



(a) (b)

Figure B.6: 2D histograms of (a) $m_Y \times m_X$ and (b) $m_h \times m_X$ showing hot spots at (3.7 TeV, 80 GeV) in $m_Y \times m_X$ and (80 GeV, 80 GeV) in $m_h \times m_X$. These values correspond to the individual excesses observed in each of these mass variables.

B.3.2 Dependence on the analysis region selections

The observed excess is not strongly dependent on the anomaly score. Figure B.7 shows that in a region inclusive of the anomaly score selection, the same feature at 3.7 TeV at 3.1σ (4.2σ local) significance is observed. This provides evidence that the excess is not an artifact of the novel VRNN implemented in this analysis, from which the anomaly score is derived.

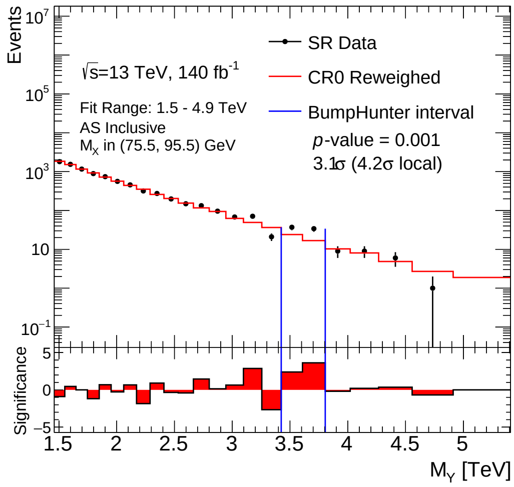
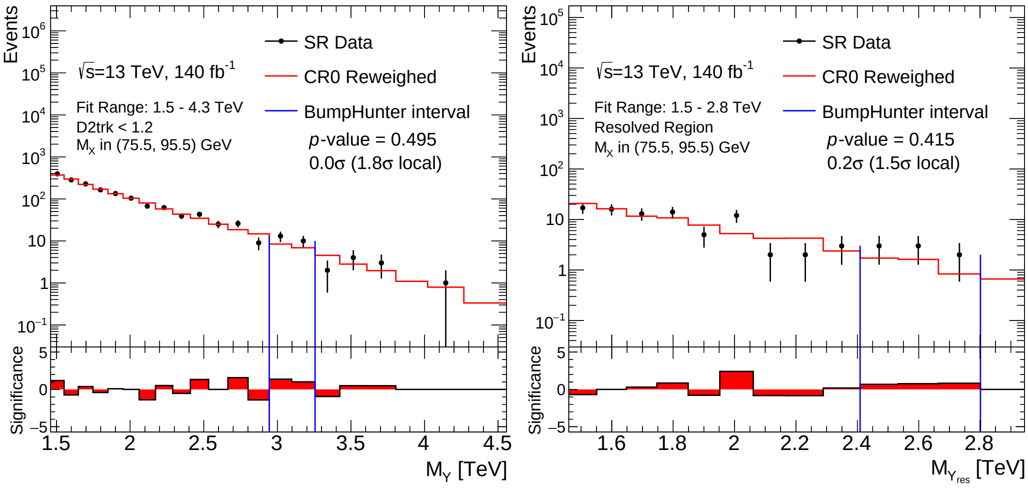


Figure B.7: BUMPHUNTER fit of m_Y inclusively of the anomaly score selection in the m_X bin of (75.5, 95.5) GeV. We observe a deviation of data from background at 3.1σ (4.2σ local) significance in the same location as in the anomaly score selected region.

The excess is not observed in either of the nominal analysis regions. This is shown on figure B.9, where the BUMPHUNTER fits in the merged and resolved regions show no significant deviation from the background. The BUMPHUNTER intervals determined from these fits are also not centered at the 3.7 TeV value of the excess.

B.3.3 Contribution from binning artifacts

The analysis bins for m_Y and m_X are chosen in a way to account for statistics in the tails of the m_Y distribution by widening the bins in the tail. the possibility that the procedure performed could be enhancing a series of fluctuations in the m_Y tail and generating a false excess. Figure B.9 shows two BUMPHUNTER fits in $m_X$bins determined from the mass resolution, with $m_Y$bins deliberately chosen very finely to dissipate a potential fluctuation. We conclude that the excess observed is resistant to binning changes, as we observe a 3.1σ (2.9σ local) and a 3.4σ (3.4σ local) significance excesses in m_X bins of (69.5, 89.5) and (74.0, 94.5) at 3.7 TeV, respectively.



(a) (b)

Figure B.8: BumpHunter fits in the Two-prong merged (a) and resolved (b) regions in the m_X bin of (75.5, 95.5). No significant deviation from the background is observed.

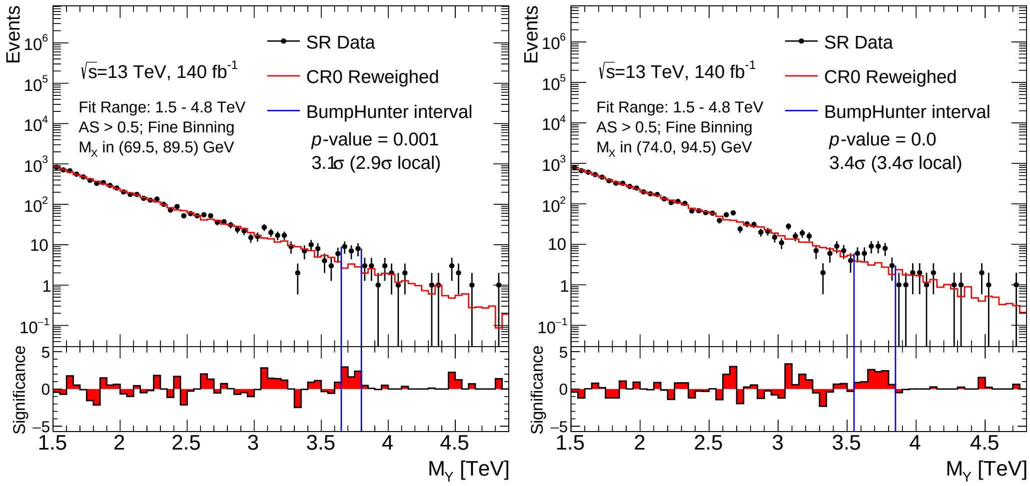
B.3.4 Contribution from detector artifacts

The possibility of a detector artifact coalescing to generate an excess in $m_Y$has been considered. The excess is centered at 3.7 TeV in m_Y and at 80 GeV in m_X, so the windows were restricted around the BumpHunter intervals of (3.4, 3.9) TeV and (70, 90) GeV in m_Y, m_X respectively and plotted both detector and kinematic variables from the events comprising the excess.

In the figures of this subsection, “SR” will represent the signal region excess events while “CR0” will represent the control region without being reweighed by the DNN weights.

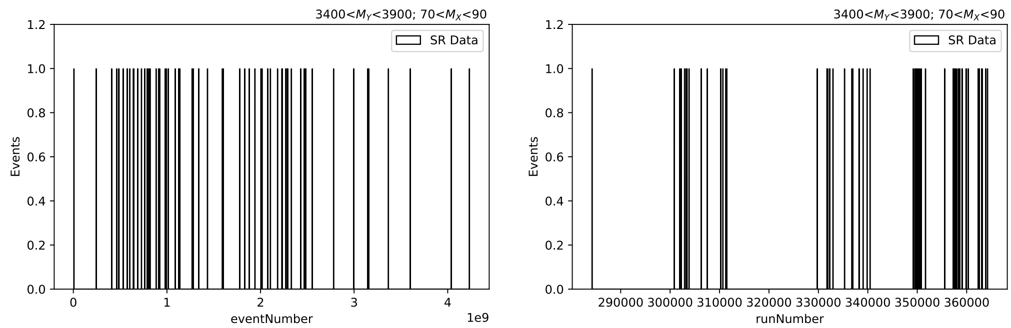
Figure B.10 shows that all the event numbers and run numbers of the excess events are unique. The run numbers are clustered towards more recent runs, though figure B.11 shows that no pile-up effect is discernible between the excess events and those in the control region. Figure B.12 shows the distribution of η, ϕ, $\Delta\eta$, and $\Delta\phi$ for both the X and Higgs candidates in the excess events. The majority of the events are produced back-to-back and lie on the endcaps, though no abnormal feature signaling a detector problem is observed.

Figure B.13 shows the p_T and E distribution of the events, with no trend between the SR and CR0 observed.



(a) (b)

Figure B.9: BUMPHUNTER fits in m_X bins of (69.5, 89.5) (a) and (74.0, 94.5) (b) as determined from the signal mass resolution, with m_Y bins chosen to be four times as fine than the nominal. We observe a 3.1σ (2.9σ local) and a 3.4σ (3.4σ local) significance excesses in (a) (69.5, 89.5) and (b) (74.0, 94.5), respectively, at 3.7 TeV.



(a) (b)

Figure B.10: Event numbers (a) and run numbers (b) for the events comprising the excess. No duplicate events or multiple events coming from the same run are observed.

B.3.5 Characterization of jet mass shapes

Figure B.5 shows a corresponding excess in m_X at 80 GeV. The observed shape of this distribution has been studies to understand whether it matches that of large-R

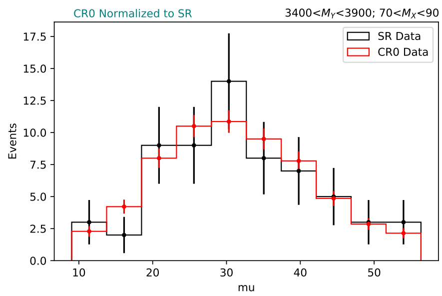
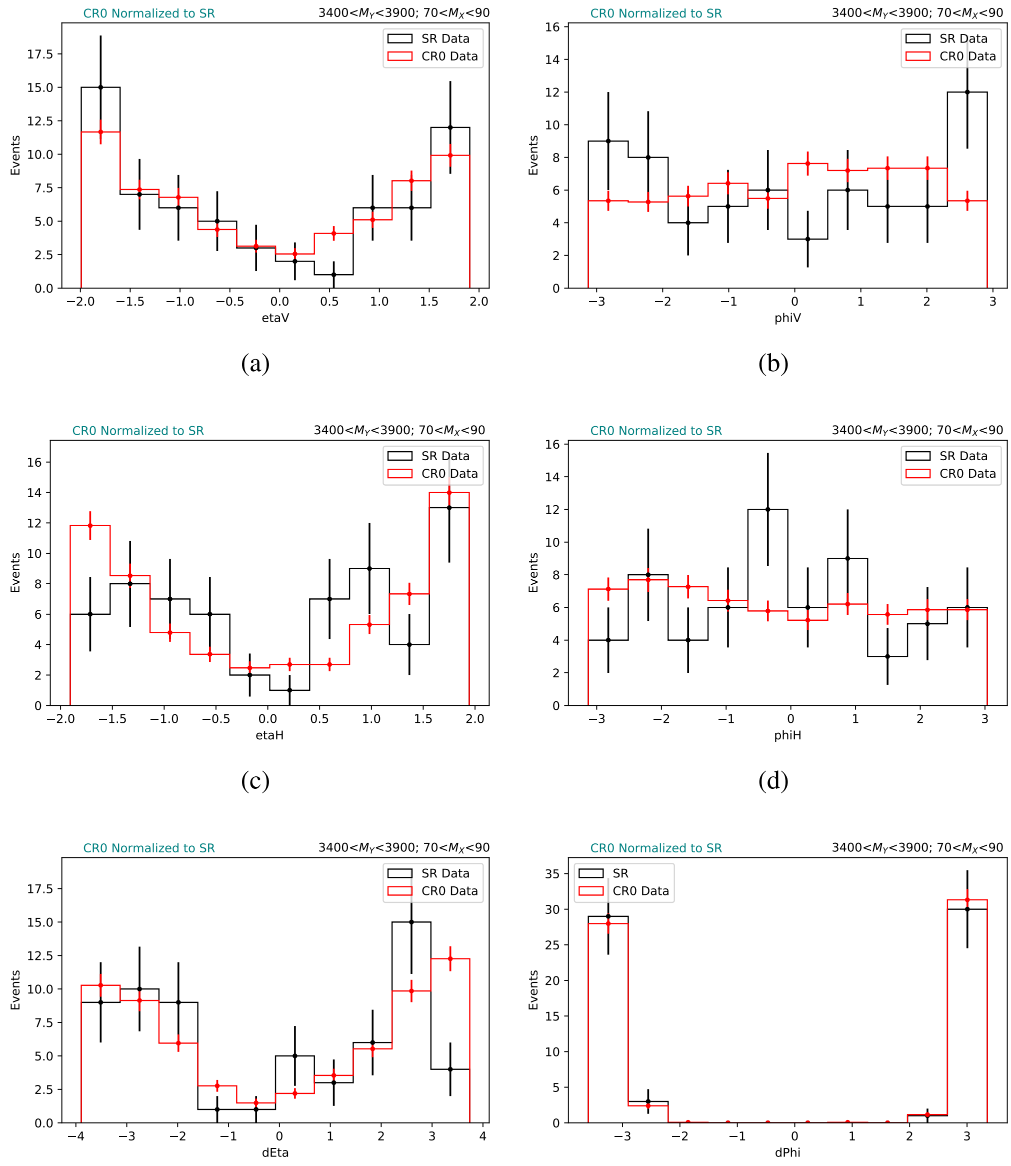


Figure B.11: Mean number of interactions per bunch crossing $\langle\mu\rangle$ for the events in the excess. No discernble deviation from control region data is observed.

jet by overlaying a $Y \to XH$signal MC with $m_X$at 80 GeV, and W/Z+jets MC to also rule out a potential contamination from W/Z events.
In Figure B.14 can be observed that the excess in m_X is too narrow to correspond to a large-R jet mass shape, or to either the $Y \to XH$or W/Z+jets events.

A contribution from W/Z events can be ruled out by noting on figure B.15, that the excess events do not show a two-pronged substructure.
From figure B.16, it is visible that from the X_{bb} Tagger probabilities, the X candidate is very QCD-like, while the Higgs candidate is very Higgs-like. A Z contribution in the Higgs candidate is possible, though ruled out, as no high bin at 90 GeV is seen in the m_H distribution in figure B.5. Some events in the analysis also contain a third jet, though it can be noted here that out of 63 events in the excess, only 10 contain a third jet. This proportion is similar in the control region, and so no abnormal behavior is observed.



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

Figure B.12: Excess events' η and ϕ for the X candidate (top) and Higgs candidates (middle), and $\Delta\eta$ and $\Delta\phi$ (bottom). The η distributions show the events are predominantly in the endcaps, and $\Delta\phi$ show they are back-to-back. No systematic difference between SR and CR0 is observed.

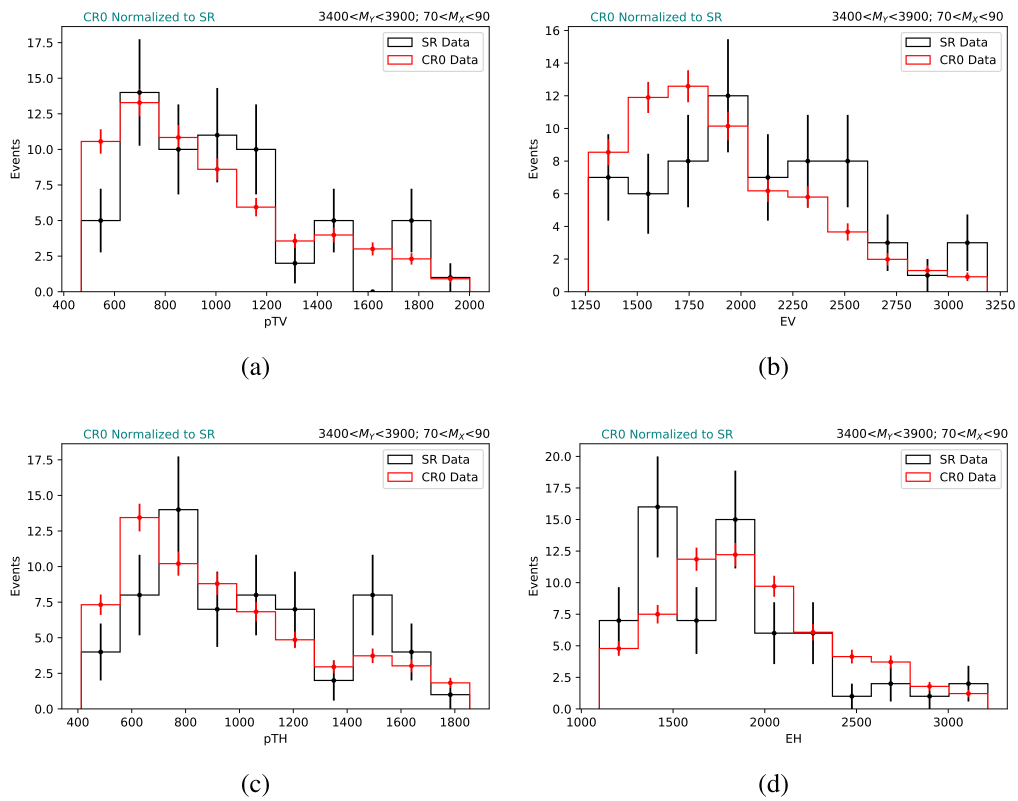


Figure B.13: Excess events' p_T and E for the X candidate (top) and Higgs candidates (bottom). No systematic trends between SR and CR0 are observed.

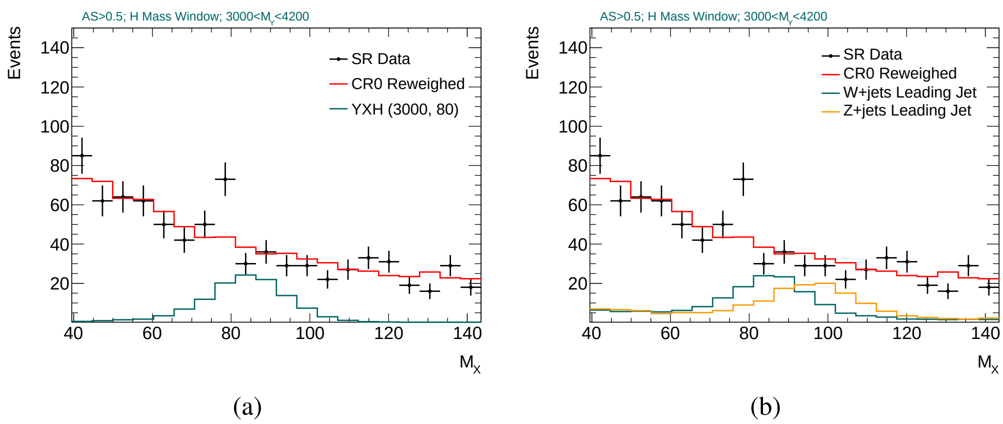
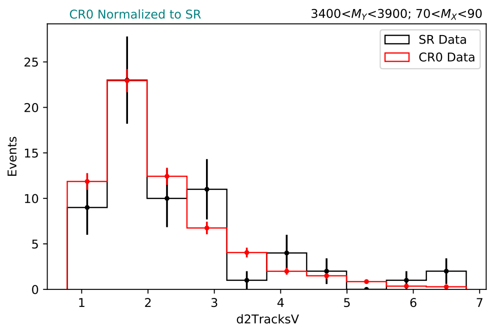
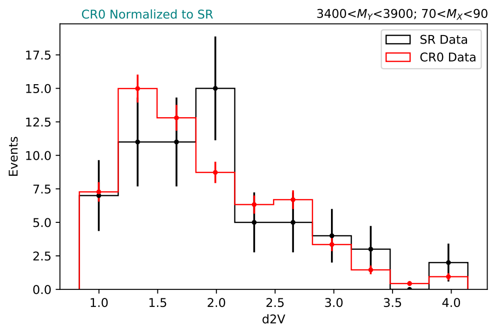


Figure B.14: Overlay of a $Y \to XH$sample MC (a) and of W/Z+jets MC (b) over the excess events in m_X. The width of the excess distribution is too narrow to correspond to either of these samples



(a)



(b)

Figure B.15: D_2^{trk}(left) and D2 (right) distribution for the events in the excess. Most of the events lie above the $D_2^{trk} < 1.2$ threshold of the merged exclusion region and these distributions do not correspond well to a two-pronged topology.

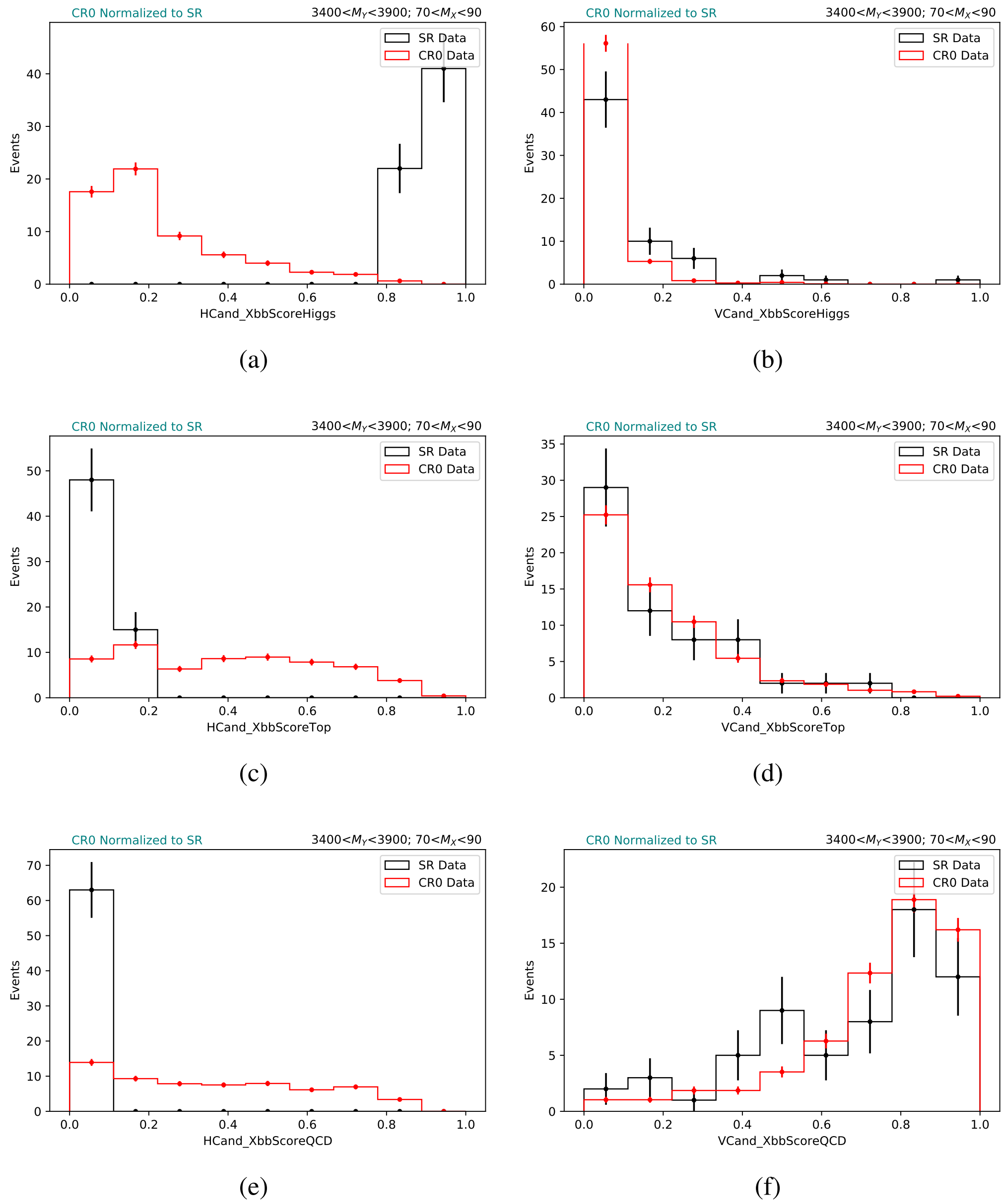


Figure B.16: X_{bb} Tagger Higgs (top), Top (middle), and QCD (bottom) probabilities for the X and Higgs candidates for the events in the excess. We observe that the Higgs candidate is very Higgs-like, while the X candidate is very QCD-like.

Bibliography

[1] the apv25 0: 25 mm cmos read out chip for the cms tracker, pre-print ic.

[2] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.

[3] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1997.

[4] Expected electron performance in the atlas experiment. Technical Report ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr 2011.

[5] Improved electron reconstruction in atlas using the gaussian sum filter-based model for bremsstrahlung. Technical Report ATLAS-CONF-2012-047, CERN, Geneva, 5 2012.

[6] Efficiency corrections for a tagger for boosted $H \to b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical report, CERN, Geneva, Aug 2021. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-035.

[7] Anomaly detection search for new resonances decaying into a Higgs boson and a generic new particle X in hadronic final states using $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. Technical report, CERN, Geneva, 2022. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2022-045.

[8] Morad Aaboud, G Aad, B Abbott, J Abdallah, Baptiste Abeloos, Syed Haide Abidi, OS AbouZeid, NL Abraham, Halina Abramowicz, Henso Abreu, et al. Jet reconstruction and performance using particle flow with the atlas detector. *The European Physical Journal C*, 77(7):1–47, 2017.

[9] Morad Aaboud, Georges Aad, Brad Abbott, B Abeloos, DK Abhayasinghe, SH Abidi, OS AbouZeid, NL Abraham, H Abramowicz, H Abreu, et al. Observation of higgs boson production in association with a top quark pair at the lhc with the atlas detector. *Physics Letters B*, 784:173–191, 2018.

[10] Morad Aaboud, Georges Aad, Brad Abbott, B Abeloos, DK Abhayasinghe, SH Abidi, OS AbouZeid, NL Abraham, H Abramowicz, H Abreu, et al. Observation of h$\rightarrow$ bb decays and vh production with the atlas detector. *Physics Letters B*, 786:59–86, 2018.

[11] Georges Aad, B Abbott, DC Abbott, A Abed Abud, K Abeling, DK Abhayasinghe, SH Abidi, OS AbouZeid, NL Abraham, H Abramowicz, et al. Muon reconstruction and identification efficiency in atlas using the full run 2 pp collision data set at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 81(7):1–44, 2021.

[12] Georges Aad, B Abbott, DC Abbott, A Abed Abud, K Abeling, DK Abhayasinghe, SH Abidi, H Abramowicz, H Abreu, Y Abulaiti, et al. Search for resonant pair production of higgs bosons in the b b¯ b b¯ final state using p p collisions at s= 13 tev with the atlas detector. *Physical Review D*, 105(9):092002, 2022.

[13] Georges Aad, Brad Abbott, Dale Charles Abbott, A Abed Abud, Kira Abeling, Deshan Kavishka Abhayasinghe, Syed Haider Abidi, OS AbouZeid, Nadine L Abraham, Halina Abramowicz, et al. Atlas b-jet identification performance and efficiency measurement with $t\bar{t}t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ tev. *The European physical journal C*, 79(11):1–36, 2019.

[14] Georges Aad, A Aggarwal, S Caron, L Colasurdo, V Fabiani, CA Gottardo, N Ilic, AC Konig, L Pedraza Diaz, and JFP Schouwenberg. Electron and photon performance measurements with the atlas detector using the 2015–2017 lhc proton-proton collision data. 2019.

[15] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale

machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[16] E Abat, JM Abdallah, TN Addy, P Adragna, M Aharrouche, A Ahmad, TPA Akesson, M Aleksa, C Alexa, K Anderson, et al. Combined performance studies for electrons at the 2004 atlas combined test-beam. *Journal of Instrumentation*, 5(11):P11006, 2010.

[17] B. Abbott, J. Albert, F. Alberti, M. Alex, G. Alimonti, S. Alkire, P. Allport, S. Altenheiner, L.S. Ancu, E. Anderssen, and et al. Production and integration of the atlas insertable b-layer. *Journal of Instrumentation*, 13(05):T05008–T05008, May 2018.

[18] J Agarwala, MG Alviggi, M Antonelli, F Anulli, C Arcangeletti, S Auricchio, P Bagnaia, S Bariani, A Baroncelli, I Bashta, et al. Construction and test of the sm1 type micromegas chambers for the upgrade of the atlas forward muon spectrometer. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, page 167285, 2022.

[19] Sea Agostinelli, John Allison, K al Amako, John Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, G 2 Barrand, et al. Geant4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[20] M Aharrouche, C Adam-Bourdarios, M Aleksa, D Banfi, D Benchekroun, K Benslama, M Boonekamp, T Carli, L Carminati, H Chen, et al. Measurement of the response of the atlas liquid argon barrel calorimeter to electrons at the 2004 combined test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 614(3):400–432, 2010.

[21] M. Aharrouche et al. Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam. *Nucl. Instrum. Meth. A*, 614:400–432, 2010.

[22] T Alexopoulos, AA Altintas, M Alviggi, M Arik, SA Cetin, V Chernyatine, E Cheu, D Della Volpe, M Dris, D Fassouliotis, et al. Development of large size micromegas detector for the upgrade of the atlas muon system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 617(1-3):161–165, 2010.

[23] T Alexopoulos, M Bianco, M Biglietti, C Bini, M Byszewski, G Iakovidis, P Iengo, M Iodice, E Karentzos, S Leontsinis, et al. Performance studies of resistive-strip bulk micromegas detectors in view of the atlas new small wheel upgrade. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 937:125–140, 2019.

[24] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box. *Journal of High Energy Physics*, 2010(6):43, 2010.

[25] Johan Alwall, R Frederix, S Frixione, V Hirschi, Fabio Maltoni, Olivier Mattelaer, H-S Shao, T Stelzer, P Torrielli, and M Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7):79, 2014.

[26] G Arnison, A Astbury, B Aubert, C Bacci, G Bauer, A Bezaguet, R Böck, TJV Bowcock, M Calvetti, T Carroll, et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$= 540 gev. *Physics letters B*, 122(1):103–116, 1983.

[27] G Arnison, A Astbury, B Aubert, C Bacci, G Bauer, A Bezaguet, R Böck, TJV Bowcock, M Calvetti, P Catz, et al. Experimental observation of lepton pairs of invariant mass around 95 gev/c^2 at the cern sps collider. *Physics Letters B*, 126(5):398–410, 1983.

[28] ATLAS Collaboration. Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter. Technical report, CERN, Geneva, Apr 2007.

[29] ATLAS Collaboration. The silicon microstrip sensors of the atlas semiconductor tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 578(1):98–118, 2007.

[30] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. Technical report, CERN, Geneva, Nov 2014.

[31] ATLAS Collaboration. Topological cell clustering in the atlas calorimeters and its performance in lhc run 1. *The European Physical Journal C*, 77(7), Jul 2017.

[32] ATLAS Collaboration. A search for resonances decaying into a higgs boson and a new particle x in the xh→qqbb final state with the atlas detector. *Physics Letters B*, 779:24–45, Apr 2018.

[33] ATLAS Collaboration. Electron and photon performance measurements with the atlas detector using the 2015–2017 lhc proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006–P12006, Dec 2019.

[34] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical report, CERN, Geneva, Jun 2019.

[35] ATLAS Collaboration. Search for diboson resonances in hadronic final states in 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 09:091, 2019. [Erratum: JHEP 06, 042 (2020)].

[36] ATLAS Collaboration. Dijet Resonance Search with Weak Supervision Using $\sqrt{s} = 13\,\text{TeV}$ pp collisions in the ATLAS detector. *Phys. Rev. Lett.*, 125:131801, 2020.

[37] ATLAS Collaboration. Search for resonances decaying into a weak vector boson and a Higgs boson in the fully hadronic final state produced in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector. *Phys. Rev. D*, 102:112008, 2020.

[38] ATLAS Collaboration. Efficiency corrections for a tagger for boosted $H \to b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical report, CERN, Geneva, Aug 2021.

[39] G. Avoni, M. Bruschi, G. Cabras, D. Caforio, N. Dehghanian, A. Floderus, B. Giacobbe, F. Giannuzzi, F. Giorgi, P. Grafström, V. Hedberg, F. Lasagni Manghi, S. Meneghini, J. Pinfold, E. Richards, C. Sbarra, N. Semprini Cesari, A. Sbrizzi, R. Soluk, G. Ucchielli, S. Valentinetti, O. Viazlo, M. Villa, C. Vittori, R. Vuillermet, and A. Zoccoli. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *Journal of Instrumentation*, 13(07):P07017–P07017, jul 2018.

[40] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke. Interpolation between multi-dimensional histograms using a new non-linear moment morphing method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 771:39–48, Jan 2015.

[41] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology. *Nucl. Phys. B*, 849:296–363, 2011.

[42] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(1-2):119–126, Jan 2008.

[43] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[44] M Capeans, G Darbo, K Einsweiler, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical report, Sep 2010.

[45] Ta-Pei Cheng and Ling-Fong Li. *Gauge theory of elementary particle physics*. Oxford university press, 1994.

[46] François Chollet et al. Keras. https://keras.io, 2015.

[47] Georgios Choudalakis. On hypothesis testing, trials factor, hypertests and the bumphunter, 2011.

[48] Junyoung Chung, Kyle Kastner, Laurent Dinh, Kratarth Goel, Aaron Courville, and Yoshua Bengio. A recurrent latent variable model for sequential data, 2015.

[49] ATLAS Collaboration. Electron performance measurements with the atlas detector using the 2010 lhc proton-proton collision data. *The European Physical Journal C*, 72(3):1909, 2012.

[50] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012.

[51] ATLAS Collaboration. Jet energy measurement with the atlas detector in proton-proton collisions at $\sqrt{s} = 7\ tev$. *The European Physical Journal C*, 73(3):2304, 2013.

[52] ATLAS Collaboration. Topological cell clustering in the atlas calorimeters and its performance in lhc run 1. *The European Physical Journal C*, 77(7):490, 2017.

[53] ATLAS Collaboration. Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 lhc proton–proton collision data at $\sqrt{s} = 13ev$. *The European Physical Journal C*, 79(8), Aug 2019.

[54] ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ tev with the atlas detector. Technical report, Jul 2020.

[55] ATLAS Collaboration. Measurements of the higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s}$ = 13 tev, 2020.

[56] ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions at s= 13 tev with the atlas detector. *Eur. Phys. J. C*, 81:689, 2021.

[57] ATLAS collaboration et al. Jet mass reconstruction with the atlas detector in early run 2 data,(2016). *Cited on*, page 64.

[58] Atlas Collaboration et al. ATLAS Pythia 8 tunes to 7 TeV data. Technical report, LHC/ATLAS Experiment, 2014.

[59] ATLAS collaboration et al. Monte carlo calibration and combination of in-situ measurements of jet energy scale, jet energy resolution and jet mass in atlas. ATLAS-CONF-2015-037, 2015.

[60] ATLAS collaboration et al. Improving jet substructure performance in atlas using track-caloclusters. Technical report, ATL-PHYS-PUB-2017-015, 2017.

[61] ATLAS Collaboration et al. Aw/z-boson tagger using track-calocluster jets with atlas. Technical report, ATLAS-PHYS-PUB-2020-008, 2020.

[62] ATLAS Collaboration et al. Identification of boosted higgs bosons decaying into bb with neural networks and variable radius subjets in atlas, 2020.

[63] Atlas Collaboration et al. Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{\{s\}}$= 13 tev with the atlas detector. *arXiv preprint arXiv:2007.02645*, 2020.

[64] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, 2012.

[65] NA48 Collaboration. A new measurement of direct cp violation in two pion decays of the neutral kaon. *Physics Letters B*, 465(1-4):335–348, 1999.

[66] The ATLAS collaboration. Operation of the ATLAS trigger system in run 2. *Journal of Instrumentation*, 15(10):P10004–P10004, oct 2020.

[67] UA2 Collaboration. Evidence for $z^0 \to e^+e^-$ at the cern pp collider. *Physics Letters B*, 129(1-2):130–140, 1983.

[68] UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern $\bar{p}p$ collider. *Phys. Lett. B*, 122(CERN-EP-83-25):476–485, 1983.

[69] Thijs G Cornelissen, M Elsing, I Gavrilenko, JF Laporte, W Liebig, M Limper, K Nikolopoulos, A Poppleton, and A Salzburger. The global χ^2 track fitter in atlas. In *Journal of Physics: Conference Series*, volume 119, page 032013. IOP Publishing, 2008.

[70] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1–19, 2011.

[71] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), Feb 2011.

[72] Kyle S. Cranmer. Kernel estimation in high-energy physics. *Comput. Phys. Commun.*, 136:198–207, 2001.

[73] Csaba Csáki. The minimal supersymmetric standard model. *Modern Physics Letters A*, 11(08):599–613, Mar 1996.

[74] Francesco Armando Di Bello. Optimisation of the atlas b-tagging algorithms for the 2017-2018 lhc data-taking. Technical report, ATL-COM-PHYS-2017-1472, 2017.

[75] John Ellis. Higgs Physics. pages 117–168. 52 p, 2015. 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report.

[76] François Englert and Robert Brout. Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters*, 13(9):321, 1964.

[77] Rudolf Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2-3):444–450, 1987.

[78] Katherine Garrett and Gintaras Duda. Dark matter: A primer. *Advances in Astronomy*, 2011:1–22, 2011.

[79] Yannis Giomataris, Ph Rebourgeard, Jean Pierre Robert, and Georges Charpak. Micromegas: a high-granularity position-sensitive gaseous detector for high particle-flux environments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 376(1):29–35, 1996.

[80] Sheldon L Glashow. Partial-symmetries of weak interactions. *Nuclear physics*, 22(4):579–588, 1961.

[81] T Gleisberg, S Höche, F Krauss, M Schönherr, S Schumann, F Siegert, and J Winter. Event generation with sherpa 1.1. *Journal of High Energy Physics*, 2009(02):007–007, Feb 2009.

[82] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken symmetries. *Phys. Rev.*, 127:965–970, Aug 1962.

[83] O. W. Greenberg. Spin and unitary-spin independence in a paraquark model of baryons and mesons. *Phys. Rev. Lett.*, 13:598–602, Nov 1964.

[84] M. Y. Han and Y. Nambu. Three-triplet model with double $\mathrm{SU}(3)$ symmetry. *Phys. Rev.*, 139:B1006–B1010, Aug 1965.

[85] FJ Hasert, S Kabe, W Krenz, J Von Krogh, D Lanske, J Morfin, K Schultze, H Weerts, G Bertrand-Coremans, Jean Sacton, et al. Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment. *Nuclear Physics B*, 73(1):1–22, 1974.

[86] Peter W Higgs. Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13(16):508, 1964.

[87] George Iakovidis. Vmm-an asic for micropattern detectors. In *EPJ Web of Conferences*, volume 174, page 07001. EDP Sciences, 2018.

[88] A. Kahn, J. Gonski, I. Ochoa, D. Williams, and G. Brooijmans. Anomalous jet identification via sequence modeling. *Journal of Instrumentation*, 16(08):P08012, aug 2021.

[89] Takafumi Kanamori, Shohei Hido, and Masashi Sugiyama. A least-squares approach to direct importance estimation. *The Journal of Machine Learning Research*, 10:1391–1445, 2009.

[90] T Kawamoto, S Vlachos, L Levinson, C Amelung, G Mikenberg, L Pontecorvo, D Lellouch, J Dubbert, C Dallapiccola, R Richter, et al. New small wheel technical design report. Technical report, 2013.

[91] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jets with variable r. *Journal of High Energy Physics*, 2009(06):059, 2009.

[92] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet trimming. *Journal of High Energy Physics*, 2010(2):84, 2010.

[93] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter clustering algorithms: Description and performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, 4 2008.

[94] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. Energy correlation functions for jet substructure. *Journal of High Energy Physics*, 2013(6), Jun 2013.

[95] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. Energy correlation functions for jet substructure. *Journal of High Energy Physics*, 2013(6), jun 2013.

[96] Hoang Ngoc Long. Su (3) l u (1) n model for right-handed neutrino neutral currents. *Physical Review D*, 54(7):4691, 1996.

[97] Sorin Martoiu, Hans Muller, and J Toledo. Front-end electronics for the scalable readout system of rd51. In *2011 IEEE Nuclear Science Symposium Conference Record*, pages 2036–2038. IEEE, 2011.

[98] Simone Marzani, Gregory Soyez, and Michael Spannowsky. Looking inside jets. *Lecture Notes in Physics*, 2019.

[99] Start-up Trigger Menu. Initial performance assessment of the atlas trigger using run-2 data, 2015.

[100] Umberto Michelucci and Francesca Venturini. Estimating neural network's performance with bootstrap: A tutorial. *Machine Learning and Knowledge Extraction*, 3(2):357–373, 2021.

[101] George V Moustakides and Kalliopi Basioti. Training neural networks for likelihood/density ratio estimation. *arXiv preprint arXiv:1911.00405*, 2019.

[102] Duccio Pappadopulo, Andrea Thamm, Riccardo Torre, and Andrea Wulzer. Heavy vector triplets: bridging theory and data. *Journal of High Energy Physics*, 2014(9):60, 2014.

[103] Giovanni Punzi. Sensitivity of searches for new signals and its optimization. *eConf*, C030908:MODT002, 2003.

[104] Alexander L Read. Presentation of search results: the cl_s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[105] A Salam. N. svartholm, ed. elementary particle physics: Relativistic groups and analyticity. In *Eighth Nobel Symposium. Stockholm: Almquvist and Wiksell*, page 367, 1968.

[106] Andreas Salzburger. Track simulation and reconstruction in the atlas experiment, 2008.

[107] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[108] Gary Steigman, Basudeb Dasgupta, and John F. Beacom. Precise relic WIMP abundance and its impact on searches for dark matter annihilation. *Physical Review D*, 86(2), jul 2012.

[109] Jesse Thaler and Ken Van Tilburg. Identifying boosted objects with n-subjettiness. *Journal of High Energy Physics*, 2011(3), Mar 2011.

[110] A Vogel. ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates. Technical report, CERN, Geneva, Apr 2013.

[111] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.