Search for new physics using diboson topologies: from the direct to the indirect quest

Reina Camacho Toro

UNIVERSITÉ PARIS CITÉ

HABILITATION A DIRIGER DES RECHERCHES
Spécialité : "Physique des particules"

Search for new physics using diboson topologies: from the direct to the indirect quest

---

Présenté et soutenue publiquement par
Dr. Reina CAMACHO TORO
Docteur de l'Université Blaise Pascal (UBP) et chercheuse CNRS
le 23 mai 2023

Jury composé de
Maria Teresa DOVA, Universidad Nacional de La Plata
Marie-Helène GENEST, LPSC-Grenoble
Gautier HAMEL de MONCHENAULT, CEA-Saclay
José OCARIZ, Université Paris Cité
Verónica SANZ, University of Sussex/IFIC-Valencia

Document submitted in fulfillment of the requirements
for the degree of Habilitation à diriger des recherches

in the

Laboratoire de Physique Nucléaire et de Hautes Energies - LPNHE
Université Paris Cité
École Doctorale STEP'UP

August 21, 2023

Contents

Introduction

In spite of the dazzling success of the Standard Model (SM), completed by the discovery in 2012 of the Higgs boson by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) there are many reasons to believe that it is not a complete theory of particle physics. Many open questions remain, such as: what is dark matter? what happened to the antimatter after the big bang? why are there three generations of quarks and leptons with such a different mass scale? the hierarchy problem and more.

There are many approaches to try to answer these questions, ranging from dedicated experiments for dark matter, antimatter, among others to the giant experiments at the LHC. We, particle physicists working at the LHC, really hope new information from experiments at the LHC will help us to find more of these missing pieces. In this context two complementary approaches have been defined in the quest for physics beyond the SM (BSM): direct searches for new physics and new particles and indirect searches through the use of constraints mainly from flavour physics observables, cosmological data and dark matter relic density, electroweak and Higgs precision tests and anomalous magnetic moment of the muon. In the last 13 years, as a member of the ATLAS collaboration, I have had the opportunity to explore both approaches, using the diboson process as the candle to shred some light on these open questions. This document will summarise this work, what I have learned in this professional journey and hopefully will also provide some discussion about possible ways to further exploit the potential of the LHC in the years to come for new physics searches.

My personal quest for BSM physics is not recent. Already during my PhD (2009-2012) at the former Université Blaise Pascal (UBP) in Clermont Ferrand (currently Université Clermont Auvergne) I worked on the search for heavy particles, in the TeV range mass scale, decaying to pairs of top-quarks [11, 163]. This was one of the first analysis in the ATLAS collaboration exploring the physics of boosted topologies (i.e. due to the high Lorentz boost two or three of the partons from the top quark decay are reconstructed as a single object).

At the beginning of 2013 I joined the ATLAS group at the Université de Genève (UniGe) in Geneva, Switzerland as a postdoctoral researcher. During my two years there I focused on the search for high mass diboson resonances (WW, ZZ and WZ) decaying fully hadronically. This search above was initially done with LHC Run-1 data at centre-of-mass energies of $\sqrt{s} = 8$ TeV, collected between 2009 and 2012. No significant deviation from the background was found with the data from Run-1 but a small 2.5σ excess was observed around 2 TeV and similar excesses had been observed in other diboson searches performed by CMS at that point in time.

The year 2015 arrived with the start of the LHC Run-2 and the chance to join the University of Chicago as a postdoctoral researcher. The start of the LHC Run-2, with an increased centre-of mass energy of $\sqrt{s} = 13$ TeV, opened great possibilities for new phenomena. Unfortunately the excess observed in the high mass diboson resonances decaying fully hadronically analysis was not confirmed by the 2015 and 2016 dataset but the mass limits were greatly improved. During those years I contributed to the first analysis looking for WH and ZH resonances decaying fully hadronically, as well as to the first global combination of several high mass resonance searches decaying to diboson and leptons. Between 2016 and 2017 I led the ATLAS Diboson/Multilepton Exotics subgroup.

In 2018 I joined the ATLAS group at Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE) as a researcher of the French National Centre for Scientific Research (CNRS). Participating still in the global combination of several high mass resonance searches decaying to diboson and leptons, I also started using the diboson topologies for an indirect search for new physics. Building on my expertise with diboson resonance searches I contributed to the first ATLAS measurement of the vector boson associated production of Higgs (VH) decaying to b-quarks using boosted topologies.

I strongly believe that for the work of an experimental particle physicist to be complete and integral it should include three main ingredients: data analysis, detector performance and instrumentation work. My detector performance involvement has been mainly related to jet calibration and jet tagging since my PhD. Quarks and gluons, and their experimental signature known as jets, are abundantly produced in these high-energy collisions and they are present in several important physics processes at the LHC and at future colliders. To maximise the LHC potential for discovery and precision measurements one must be able to increase our understanding of jets and with that the understanding of QCD. Among other things, I was responsible for the algorithms recommendations to be used for the identification of boosted W/Z bosons jets at the beginning of Run-2. Between 2019 and 2021 I coordinated the flavour tagging calibration subgroup in ATLAS which calibrated the algorithms used to identify b-jets as well as boosted bb topologies like the Higgs ($H \to b\bar{b}$). Since October 2021 I am one of the two coordinators of the ATLAS Jets and Missing Transverse Energy (jetEtmiss), a role of 2 years.

Regarding my work on instrumentation. During my post-doc at University of Geneva, I participated in the testbeam data reconstruction from a testbeam telescope designed by UniGe-ATLAS in order to test a prototype HV-CMOS chip developed for the ATLAS Inner Tracker (ITk) upgrade as part of the High Luminosity (HL)-LHC program. During my time at the University of Chicago I worked on the Global Feature EXtractor (gFEX) project. gFEX is one of the modules designed for the upgrade of the ATLAS trigger system (L1Calo) for Run-3 and will have access to the information of the whole calorimeter needed for the reconstruction of large jets from boosted objects and the calculation of global event variables such as the missing transverse energy at the trigger level. I joined the L1Calo team and did shifts in 2015 and 2016 to get to know the previous system better while coordinating the gFEX - calorimeter integration tests that took place in 2016. When I joined the LPNHE in 2018 I re-joined the ITk community in preparation for the HL-LHC phase.

Scientific and education capacity building are also part of my core activities. In 2014, with several Venezuelan colleagues, I co-founded an organisation called Centro Virtual de Altos Estudios de Altas Energias-Chapter Venezuela (CEVALE2VE). CEVALE2VE was an initiative to strengthen academic training and research in the area of High Energy Physics (HEP). Initially created for and from the Venezuelan academic community, it expanded to several Latin American countries and brought together Venezuelan, Colombian, Peruvian and Mexican scientists involved in international collaborations in the fields of HEP, detector development and computer science. CEVALE2VE in cooperation with the International Center for Theoretical Physics (ICTP) in Trieste, Italy, organised several outreach roadtrips in South America between 2016 and 2019 through the Physics without Frontiers program for South America and the BrainGain-Venezuela program.

The community strengthened by these initiatives, and other European-Latin American programs like HELEN and e-PLANET, joined efforts in the creation of the Latin-American alliance for Capacity buildiNG in Advanced physics (LA-CoNGA physics). LA-CoNGA physics is building a network of common capacities in advanced training in Latin American universities. Initially co-financed as an Erasmus+ project (Capacity Building in Higher Education) for 4 years it has succeeded in bringing together 8 universities in Colombia, Ecuador, Peru, and Venezuela to support their efforts towards internationalisation, accessibility and digitisation. It strengthens collaboration with Europe including 3 European universities in France and Germany, and several international and national research centers (including ICTP in Italy, CERN in Switzerland and CNRS, CEA in France) as well as industrial partners (the Italian instrumentation company CAEN and data science start-ups). The project has successfully piloted the creation of a Masters-level year with key subjects of study Data Science, Instrumentation and Advanced Physics, as skills highly in demand outside academia. I have the pleasure to be the deputy coordinator of LA-CONGA physics since it started in 2020. I have also actively participated in several local outreach and science communication activities like Fêtes de la Science in Paris, guided tours to CERN and ATLAS, la "Nuit de la Science" and "Expanding your Horizons" in Geneva and masterclasses hands-on particle physics. All these activities shaped my work values and mission: diversity and collaboration in science, development of early career scientist and streghtening the relationship between Latin American and Europe.

The present document describes the main subjects I have worked on after my PhD, focusing on my personal contributions. The experimental tools: the LHC and the ATLAS detector, as well as the objects reconstruction and calibration in ATLAS are briefly introduced in Chapter 1. Chapter 2 is devoted to the direct searches for hadronically decaying diboson resonances using LHC Run-1, while the results using early (2015+2016) Run-2 data and their combination with other diboson searches are discussed in Chapter 3. Chapter 4 details the measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum and the measurement of the jet mass in high transverse momentum $Z(\rightarrow b\bar{b})\gamma$ production, both examples of indirect BSM searches using diboson topologies. A brief outlook and conclusion is presented in Chapter 5. I have also included a couple of appendices presenting my teaching and supervision of students' research activities, as well as other responsibilities held in the past years.

Several subjects that I worked on are not treated at length in this document. Those include VH resonance searches [10, 95, 96], HL-LHC prospect studies for the SM $VH, H \rightarrow b\bar{b}$ [12], searches for SUSY chargino and neutralino production in final states with a Higgs boson and missing transverse momentum [89], gFEX [139], ITk, data quality and control-room shifts for the Run-2 L1Calo system, as well as my contribution to capacity building initiatives and outreach activities.

Chapter 1

ATLAS and the LHC: a jetty place

1.1 The Large Hadron Collider (LHC)

The LHC [178] is a 27 km two-ring superconducting collider installed between 45 m and 170 m below the surface in the French–Swiss border, in the vicinity of Geneva. The last step of the CERN accelerator complex, illustrated in Figure 1.1, the LHC uses superconducting magnets to bend proton beams with energy up to a design value of 7 TeV, hence the maximum achievable center-of-mass energy is $\sqrt{s} = 14$ TeV for proton-proton (pp) collisions, and heavy ions such as lead to 5 TeV per nucleon.

Collisions take place at up to 40 MHz (25 ns bunch spacing) in four different points, all instrumented with large experiments: two general purpose experiments ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider beauty). The pp collision rate divided by the interaction cross-section, known as instantaneous luminosity, exceeds 10 nb^{-1} s^{-1} or 10^{34} cm^{-2} s^{-1}. The collision center-of-mass energy as well as the integrated luminosities delivered have changed since the beginning of the data taking in 2010. Figure 1.2 show the integrated luminosities delivered in the various years of LHC operation.

At the time of writing, we are at the beginning of the data taking period known as Run-3, which started in July 2022 and is expected to last until 2025. The LHC is expected to continue its operations until 2038 and more details about the expected performance in the next runs is given in Chapter 5. This document presents results using data from Run-1 (2010-2012) and Run-2 (2015-2018).

1.2 The ATLAS detector

The ATLAS detector [116] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle[1]. ATLAS has a wide physics program aiming basically at: Standard Model precision measurements, searches for and measurements of the Higgs boson and searches for physics beyond the Standard Model. As a collaboration, ATLAS consists of more than 5500 scientists and engineers from 245 institutes in 42 countries around the world. I am honored to be part of this collaboration since 2009, exactly when the data taking started.

Like many collider experiments, ATLAS is built of several sub-detectors, configured in concentric layers around the interaction point, each one serving a different purpose. Its

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector. The positive x-axis is defined by the direction from the interaction point to the centre of the LHC ring, with the positive y-axis pointing upwards, while the beam direction defines the z-axis. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity η is defined in terms of the polar angle θ by $\eta = -\ln\tan(\theta/2)$. Rapidity is defined as $y = 0.5\ln[(E + p_z)/(E - p_z)]$ where E denotes the energy and p_z is the component of the momentum along the beam direction. The angular distance ΔR is defined as $\sqrt{(\Delta y)^2 + (\Delta\phi)^2}$.

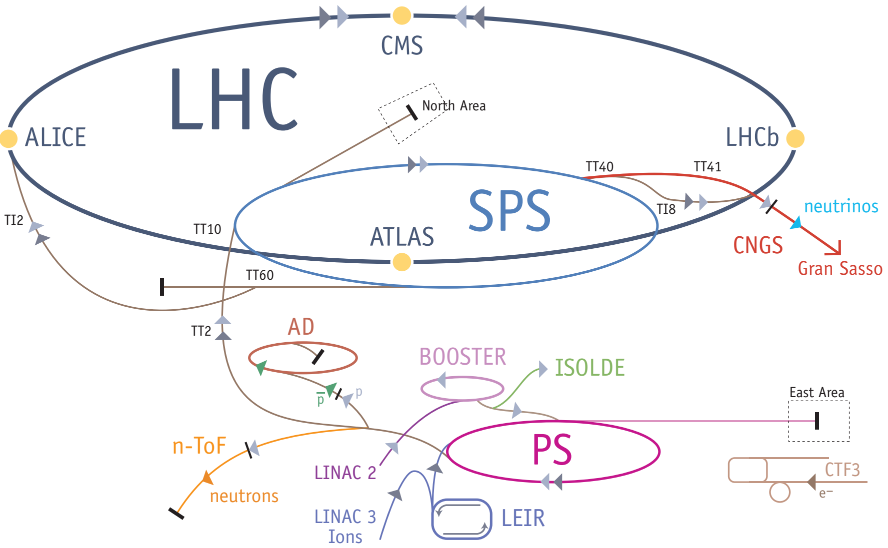


FIGURE 1.1: Illustration of the LHC injector complex. The four main LHC experiments are indicated with yellow circles.

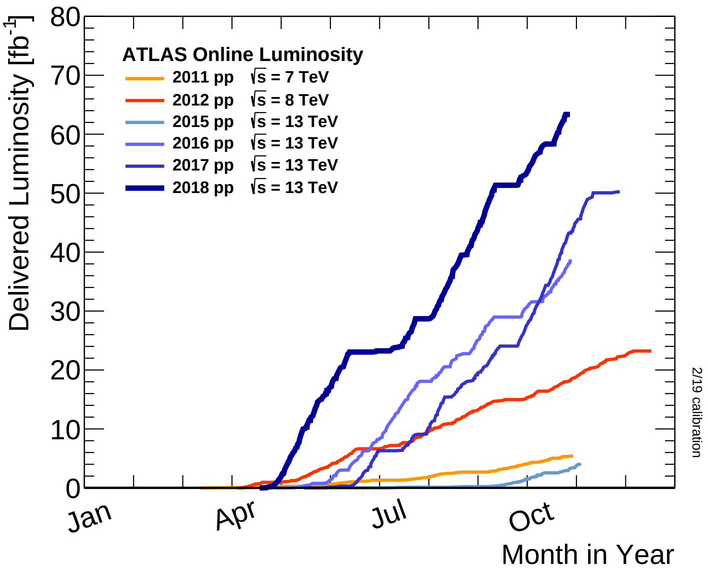


FIGURE 1.2: Cumulative integrated luminosity by year delivered to ATLAS during stable beams and for high energy p-p collisions. From Ref. [123].

layout is shown in Figure 1.3. From the interaction point outwards, the first subdetector is the inner detector (ID) which is used for tracking of charged particles. The ID covers the pseudorapidity range $|\eta| < 2.5$, has a total length of 7024 mm, an inner radius of 45 mm and an outer radius of 1150 mm, it is as close as 10 mm to the beam pipe. It consists of three tracking devices: a silicon pixel detector , a silicon microstrip detector (the Semi-Conductor Tracker SCT), and a transition radiation tracker (TRT) in the range $|\eta| < 2.0$, that provides electron identification capabilities. An inner pixel layer, the insertable B-layer [14], was added at a mean radius of 3.3 cm during the long shutdown period between Run 1 and Run 2 of the LHC. Surrounded by a thin superconducting solenoid providing a 2 T magnetic field, the ID guarantees excellent tracking capability with precise momentum determination over a wide range of momenta. It also identifies tracks from secondary vertices, permitting the efficient b-quark jets and tau-leptons identification.

The subsequent subdetector is the calorimeter system, which measures the energy of particles by total absorption. It is divided into an electromagnetic (in charge of measuring the energy and position of photons and electrons with high precision) and a hadronic component (where hadrons and other particles with enough energy to not be fully absorbed in the electromagnetic calorimeters deposit their remaining energy). Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic energy measurements with high granularity ($|\eta| < 3.2$). In the $|\eta| < 1.8$ region, it is complemented by a pre-sampler LAr-based layer used to correct for fluctuations in upstream energy losses. The calorimeter itself is separated into three layers in depth. The first layer is finely segmented up to $|\eta| < 2.5$, allowing for efficient separation of single photons from two collimated photons from a $\pi^0 \rightarrow \gamma\gamma$ decays. The second layer collects of the energy deposited in the calorimeter by photon and electron showers. The third layer completes this collection for the highly energetic e and γ and checks their isolation from the hadronic calorimeter to limit the rate of misidentified jets. The hadronic calorimeter uses a steel/scintillator-tile sampling detector in the central pseudorapidity range ($|\eta| < 1.7$, known as hadronic barrel calorimeter) and a copper/LAr detector in the region $1.5 < |\eta| < 3.2$ (known as hadronic end-cap calorimeter). As for the EM calorimeter the read-out is segmented longitudinally. In this case in three layers in both barrel and extended barrel. The cell granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the first two layers and $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$ for the third one. The forward regions ($3.2 < |\eta| < 4.9$) are instrumented with copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements, respectively.

Of all known interacting particles, only muons and neutrinos are penetrating enough to reach beyond the calorimeters. Neutrinos interact only weakly and escape the detector without leaving traces. ATLAS' almost hermetic calorimetry is fundamental for the measurement of missing transverse energy and therefore infer the production of non-interacting particles like neutrinos. A tracking sub-detector is placed outside of the calorimeters to measure the muon's momentum: the muon spectrometer. In this way the detection of all the particles is achieved. Both tracking sub-detectors (the inner detector and the muon spectrometer) operate in a magnetic field, provided by a solenoidal and a toroidal magnet system respectively. The systems most relevant for the analyses presented in this document are the inner detector and the calorimeters. Additional detectors exist to monitor the luminosity and the beam position. Events are selected using a first-level trigger implemented in custom electronics, which reduces the event rate to a design value of at most 100 kHz using a subset of detector information [85]. Software algorithms with access to the full detector information are then used in the high-level trigger to yield an average recorded event rate of about 1 kHz.

1.3 Object reconstruction and calibration in ATLAS

The algorithms used to reconstruct and identify the different objects from the detector information have of course evolved over time and depend also on the needs of each analysis, as will be mentioned in the next chapters when necessary. In this section you can find a global description of the objects reconstruction and calibration in ATLAS. *A more detailed description will be provided for jets to showcase some of my contributions on this topic*.

Electrons are reconstructed from their energy deposited in a cluster of cells in the electromagnetic calorimeter, a cluster that is associated with the track left in the ID. The identification of the electrons is based on many variables like the shape of the cluster of energy

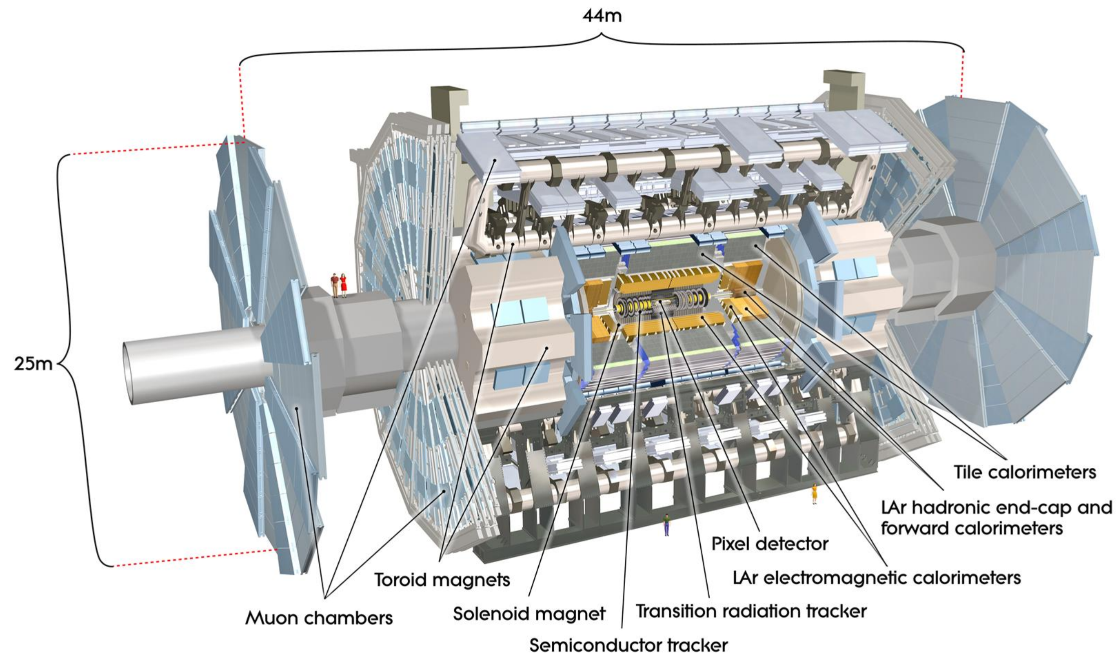


FIGURE 1.3: A detailed computer-generated image of the ATLAS detector and its subsystems. From Ref. [116].

deposits in the calorimeter, track quality and track-to-cluster matching [32, 33].

Muons are usually identified by combining a trace in the internal detector with a trace in the muon spectrometer [72]. They are required to have a minimum number of hits in each system, and must be compatible in terms of geometrical and momentum matching. It is also possible to reconstruct muons whose trace in the spectrometer is not complete but associated with a trace in the internal detector, or whose trace in the muon spectrometer points to the interaction point. Requirements on the leptons isolation are usually used at the analysis level, requiring no near-by tracks within a $p_{\mathrm{T}}-$dependent variable-size R cone around the electron.

Photon candidates are reconstructed from clusters of energy deposits in the EM calorimeter [31, 62]. The photon energy is calibrated by applying the energy scales measured with $Z \to e^+e^-$ decays. Identification requirements are applied to reduce the contamination from π^0 or other neutral hadrons decaying into photons. Requirements on the shower shape in the EM calorimeter and on the energy fraction measured in the hadronic calorimeter are used to identify photons.

For electrons, muons and photons, different identification and isolation levels are possible (from loose to tight) depending on the desired level of background rejection, knowing that an increase in purity through stricter selections leads to a decrease in identification efficiency.

The presence of very weekly interacting particles like neutrinos in some analysis topologies can be inferred from a momentum imbalance in the transverse plane. The **missing transverse momentum**, $E_{\mathrm{T}}^{\mathrm{miss}}$, is reconstructed as the negative vector sum of the momenta of leptons and jets in the event plus a 'soft term' built from additional from topo-clusters with $|\eta| < 4.9$ in Run-1 [83] and from tracks associated with the primary vertex in Run-2 [84]. Even though the track-based soft-term misses components from soft neutral particles, it was found to be beneficial, as it removes any pile-up contributions compared to the calorimeter-based variant which only uses noise suppression in the topo-clustering algorithm to avoid fake signals. The magnitude of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ is referred to as $E_{\mathrm{T}}^{\mathrm{miss}}$ in the following chapters.

1.4 A special section for jets

Jets are the experimental signatures of quarks and gluons. While the resulting bunch of particles from quarks and gluon hadronisation pass through the ATLAS detector, they produce tracks in the ID and energy deposits inside the calorimeters. Jets are key for all the different analyses that will be presented in this document. The jet reconstruction process consists in three steps: the definition of inputs, the use of a jet reconstruction algorithm to group the inputs and finally the jet calibration for jet reconstructed from detectors inputs which corrects the jet energy, momentum and mass for the effects of ATLAS calorimeters non-compensation, dead material, leakage, out of cone and other thresholds effects. Depending on the inputs and reconstruction algorithm that is used to reconstruct this collimated spray of hadrons to gain information about the initial partons, the resulting jets can be completely different objects. Finally the identification of the origin of jets, i.e. whether it is coming from a light-quark, a gluon, a $b-$jet or a hadronically decaying boson, is essential for many physics analyses. In this section I will provide a summary of the jet reconstruction process, with a focus on the calibration and jet identification algorithms where I directly contributed in the last year.

1.4.1 Jet inputs

According to the inputs used, jets can be classified as:

- Truth jets: reconstructed using stable particles [2] produced by the hadronization models of the different Monte Carlo generators, before any detector simulation. Muons and neutrinos are not taken into account. These jets are are mainly used as the reference objects for simulation-based ATLAS jet calibrations and as reference in physics studies.
- Track jets: reconstructed using charged-particle track information [29]. They are mainly used as proxies for b-jets in dense-environments (see Section 1.4.5).
- Calorimeter jets: reconstructed using noise-suppressed topological clusters (topoclusters) of calorimeter energy depositions [120].
- Jets can also be built from a combination of tracks and calibrated clusters of energy deposits in calorimeter cells. There are different ways of combine them e.g. particle flow (PFlow) [52], Track-CaloCluster (TCC) [47] or UFO [164] jets. By combining calorimeter and tracking information, one benefits from both the better energy resolution of the calorimeter at high energy and the superior angular resolution for the tracks. This is a relatively recent development in ATLAS.

  PFlow attempts to decide on a particle-by-particle basis whether the energy measurement is better performed by the tracker or calorimeter based on the relative resolution of the two detectors for particles of various energies. TCC uses the angular energy flow information from the tracker in combination with the energy measurement from the calorimeter, with a particular focus on using tracking information to resolve distinct particles which would otherwise be reconstructed by a single topocluster. And UFO combines PFlow and TCC to get a good performance across the whole jet p_{T} range.

[2] Stable particles refer to particles with a lifetime greater than 10 ps.

1.4.2 Jet reconstruction algorithms

Jet reconstruction algorithms allow to associate the inputs constituents to a jet object. Jet algorithms that are currently being used in ATLAS belong to the sequential reconstruction algorithms category, since they are infrared and collinear safe giving a stable and precise description of QCD interactions during the pp collision. Sequential recombination algorithms take inputs and combine them to form jets according to the following distance parameter:

$$d_{ij} = \min(p_{\mathrm{T},i}^{2p}, p_{\mathrm{T},j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \tag{1.1}$$

$$d_{iB} = p_{\mathrm{T},i}^{2p} \tag{1.2}$$

where ΔR_{ij} is the angular distance between the two objects i and j and R is the radius parameter of the jet algorithm which regulates the size of the jets. The free parameter of the algorithm, p, defines how the input particles are combined. The k_t algorithm [143, 177] corresponds to $p = 1$, the Cambridge-Aachen (C/A) algorithm [171] to $p = 0$ and the anti-k_t algorithm to $p = -1$ [137]. For the C/A algorithm, entities are combined purely based on their angular separation. For the k_t algorithm, soft particles are favoured to be merged first in contrast to the anti-k_t algorithm where the jet starts growing outwards from a high-p_T object. The anti-k_t algorithm is the default algorithm used in ATLAS, since due to their conical shape they are less susceptible to the underlying event and pile-up and they are also easier to calibrate. However, the irregular shapes of the k_t and C/A algorithm are useful in terms of studying the parton shower of the initial particle as the clustering process of the two algorithms reverses this process. The value used for the radius parameter R depends on the application in mind. In ATLAS jets built using a radius parameter R equal to 0.4 are referred to as "small-R" jets and they are used as proxies for single quarks or/and gluons. Jets built using a larger R, usually $R = 1.0$, are referred to as "large-R" jets and are used to reconstruct boosted particles decaying hadronically like top-quarks and bosons.

1.4.3 Jet pile-up mitigation

Since the beginning of data taking in 2009 the average number of pileup interactions per bunch crossing has increased significantly, exacerbating the pileup dependence of existing jet reconstruction procedures. In Run-1 the average of interactions per crossing was 21, 14 in 2015 at the beginning of Run-2 and 36 in 2018. For small-R jets, this dependence is mitigated using a two step method: a jet-area-based subtraction of pile-up energy method and a the residual pile-up correction dependent on the number of reconstructed primary vertices, N_{PV}, and the mean number of interactions per bunch crossing, μ [51].

For large-R jets, this dependence is usually minimized through the use of jet inputs which suppress calorimeter noise like noise-suppressed topological clusters and jet grooming algorithms (which will be described shortly). Towards the end of Run-2 this treatment was not enough. Recently some further constituent-level pileup mitigation techniques have been implemented and tested in ATLAS to further reduce the residual pileup dependence, e.g. Constituent Subtraction [127], Voronoi Subtraction [24], SoftKiller [136], and PUPPI [128].

Jet grooming algorithms generally have two uses; (i): to remove contributions from pileup , and (ii) to reveal hard substructure within jets resulting from massive particle decays by removing the soft component of the radiation. There are many grooming algorithms developed in the last years. Here we will briefly describe those will be useful in next chapters, the description is fully taken from Ref. **??**.

- Trimming [192]: Starting with constituents of jets initially reconstructed using the C/A or anti-k_t algorithm, smaller 'subjets' are reconstructed using the k_t algorithm with a radius parameter $R = R_{\text{sub}}$, and removed if they carry less than a fraction f_{cut} of the original, ungroomed, large-R jet p_{T}. For reference, the recommended trimming configuration used in ATLAS studies for the majority of Run-2 analyses was anti-k_t, $R = 1.0$, with $f_{\text{cut}} >= 5\%$ and $R_{\text{sub}} = 0.2$. This was the results of a detailed study of techniques for identifying boosted, hadronically decaying bosons to which I contributed and will be described in Chapter 2.

- Pruning [176]: The constituents of jets initially reconstructed with the C/A or anti-k_t algorithms are re-clustered with the C/A algorithm with two parameters: R_{cut} and Z_{cut}. In each pairwise clustering, the secondary constituent is discarded if it is (i) wide-angled: $\Delta R_{12} > R_{\text{cut}} \times 2M/p_{\text{T}}$, where ΔR_{12} is the angular separation of the two subjets; or (ii) soft: $f_2 < Z_{\text{cut}}$, where M is the jet mass and f_2 is the p_{T} fraction of the softer constituent with respect to the p_{T} of the pair. A configuration of the pruning algorithm was favoured by the CMS experiment for boson tagging [150], using C/A jets with $R = 0.8$ and pruning with $Z_{\text{cut}} = 10\%$ and $R_{\text{cut}} = 1$ during Run-1 and beginning on Run-2.

- Split-filtering [134]: This algorithm has two stages: the first (splitting) is based on the jet substructure, and the second (filtering) is a grooming stage to remove soft radiation. For the first stage, C/A jets are de-clustered through the clustering history of the jet. This declustering is an exact reversal of the C/A clustering procedure, and can be thought of as splitting the jet into two pieces. The momentum balance, $\sqrt{y_{12}}$, is defined as:

$$\sqrt{y_{12}} = \frac{\min(p_{\text{T}1}, p_{\text{T}2})}{m_{12}} \Delta R_{12}, \tag{1.3}$$

where $p_{\text{T}1}(p_{\text{T}1})$ is the piece with the highest (the lowest) p_{T}, and m_{12} is the invariant mass of the two pieces. The mass-drop fraction μ_{12} is the fraction of mass carried by the piece with the highest mass:

$$\mu_{12} = \frac{\max(m_1, m_2)}{m_{12}}. \tag{1.4}$$

If the requirements on the mass-drop $\mu_{12} < \mu_{\text{max}}$ and momentum balance $\sqrt{y_{12}} > \sqrt{y_{\text{min}}}$ are met then the jet is accepted and can proceed to the filtering stage. Otherwise the de-clustering procedure continues with the highest mass piece: this is now split into two pieces and the μ_{12} and $\sqrt{y_{12}}$ requirements are again checked. This process continues iteratively. In the filtering stage, the constituents of the surviving jet are reclustered with a subjet size of $R_{\text{sub}} = \min(0.3, \Delta R_{12})$ where ΔR_{12} is taken from the splitting stage. Any remaining radiation outside the three hardest subjets is discarded. This algorithm differs somewhat from pruning and trimming in that it involves both grooming and jet selection. A version of this algorithm was favoured by the ATLAS diboson resonance searches performed in Run-1 described in Chapter 2.

- Soft-drop [198]: In ATLAS soft-drop calorimeter jets are formed by the application of the soft-drop algorithm [198] to the anti-k_t $R = 1.0$ jets described above, with additional topological cluster preprocessing that is described below. The soft-drop algorithm is designed to remove soft and wide-angle radiation and also contamination from pileup. In the first step of the grooming algorithm, the anti-k_t $R = 1.0$ jets are reclustered with the Cambridge-Aachen (C/A) algorithm so that the constituents are combined purely according to their angular separation. The soft-drop algorithm then reverses the C/A algorithm clustering history and removes the softer subjet at a specific step of the C/A clustering history unless the soft-drop condition is fulfilled:

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0}\right)^{\beta} , \tag{1.5}$$

where z_{cut} and β are algorithm parameters, p_{T1} and p_{T2} are the transverse momenta of the declustered subjets at each history step, ΔR_{12} is the distance between the subjets in the (η, ϕ) space and R_0 is a threshold corresponding to the jet radius. The use of Soft-Drop in ATLAS is very recent and studies in Ref. [45] recommend the use of parameters $\beta = 0$ and $z_{\text{cut}} = 0.1$. These studies showed that soft-drop jet mass exhibits a pileup dependence with the chosen parameters and therefore a special version of pileup suppressed topological clusters are used to construct the jets that are then groomed with the soft-drop algorithm: the SoftKiller (SK) algorithm is used in conjunction with Constituent Subtraction mentioned before.

1.4.4 Jet calibration

The jet energy scale (JES) and jet mass scale (JMS) calibration restores the jet energy and mass to that of jets reconstructed at the particle level, respectively. The calibration chain is different for small-R and large-R jets and it has also varied along the years. In this section you will find a global description of the small-R jets calibration (taken from Ref. [51]) and a more detailed historical review of the large-R jets calibration in which I worked between 2013 and 2015 (largely based on Ref. [170]). Between the beginning of 2013 and 2016, I co-supervised the thesis of Chris Delitzsch at University of Geneva, her qualification task was related to the calibration work for large-R jets. The main goal in this period was to derive the calibrations and corresponding uncertainties for the legacy analyses using full Run-1 dataset, as well as the ones needed for the analyses using the first data from Run-2. Three jet collections were calibrated:

- anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ [27, 53]: the jet collection used by most of the ATLAS physics analyses in Run-1. More details about how this optimized choice was made can be found in Section 2.3.1.
- split-filtered C/A jets with $R = 1.2$ [99]: used in diboson resonance searches as described in Section 2.1
- anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$ [40, 49]: the jet collection used by ATLAS physics analyses in Run-2. The boosted vector boson identification algorithm, including the jet reconstruction and grooming algorithm, was optimised to account for the higher centre-of-mass energy and changed pile-up conditions in Run-2 as described in Section 2.3.2.

The calibrations that will be described here are essential elements of the analyses presented in Chapters 2, 3 and 4 and they were used by many other physics analyses in the ATLAS collaboration.

Jet calibration for small-R jets

The full calibration chain for small-R jets is presented in Figure 1.4. At the beginning of the small-R jets calibration chain, the pile-up corrections remove the excess energy due to pile-up interactions. These corrections consist of two components: a correction based on the jet area and transverse momentum density of the event, and a residual correction derived from MC simulation and parameterized as a function of the mean number of interactions per bunch crossing (μ) and the number of reconstructed primary vertices in the event (N_{PV}). Then, the absolute JES calibration corrects the jet so that it agrees in energy and direction with truth jets from dijet MC events. Afterwards, the global sequential calibration (GSC, derived from dijet MC events) improves the jet p_{T} resolution and associated uncertainties by removing the dependence of the reconstructed jet response on observables constructed using information from the tracking, calorimeter, and muon chamber detector systems. Finally, a residual in situ calibration is applied to correct for remaining differences between data and MC simulation. It is derived using a variety of well-measured reference objects to cover the full detector region, including photons, Z bosons, and calibrated jets, and for the first time benefits from a low-p_{T} measurement using the $E_{\mathrm{T}}^{\mathrm{miss}}$ projection fraction method for better pile-up robustness. A full description of the calibration and its corresponding uncertainties can be found in Ref. [51].

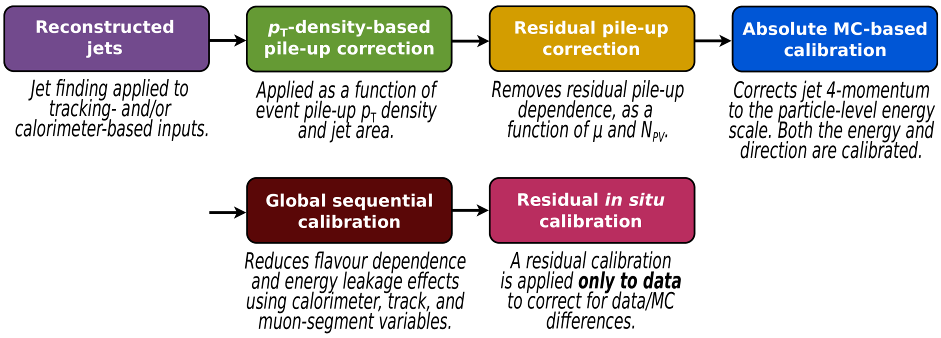


FIGURE 1.4: Stages of small-R jet energy scale calibration at the end of Run-2. Each one is applied to the four-momentum of the jet [51].

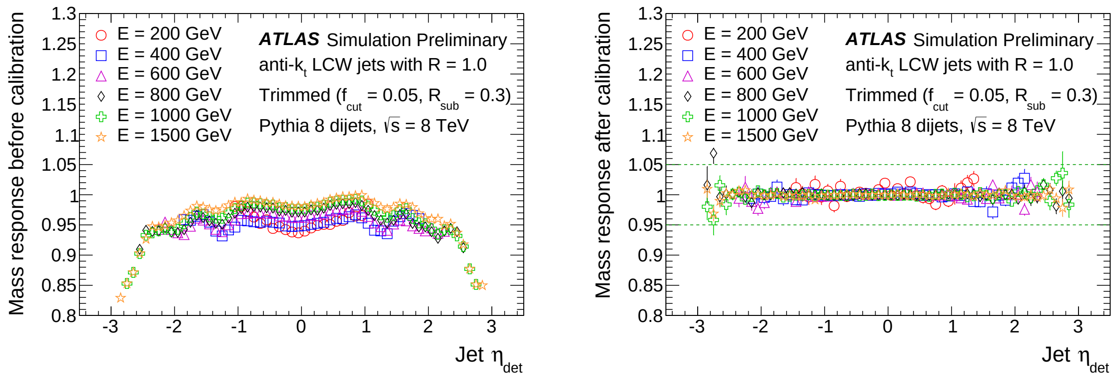
Recent results, made public in 2022 [73], exploring new calibration techniques report improvements in the jet energy scale and resolution of up to 40% and more pile-up robustness thanks to a residual correction taking into account the correlations between μ, N_{PV}, jet η and p_{T} as well as an updated version of the global sequential calibration using Neural Networks and as inputs the all of the inputs used in the GSC correction, the jet p_{T}, the jet η, μ, N_{PV}, the fraction of jet energy measured in the different calorimeter layers and the minimum number of clusters containing 90% of the jet energy.

Jet calibration for large-R jets

Back in 2013 when I started working on the large-R jets calibration the chain consisted in two absolute consecutive corrections applied to jets after the grooming. The absolute JES and JMS calibration correct the jet energy and mass to that of truth jets, respectively. The calibration factors were derived in QCD dijet MC samples but were applied to the data as well. The correction is calculated using all reconstructed groomed jets that have a matching groomed truth jet within $\Delta R = \sqrt{(\eta_{\mathrm{reco}} - \eta_{\mathrm{truth}})^2 + (\phi_{\mathrm{reco}} - \phi_{\mathrm{truth}})^2} < 0.75 \cdot R$, where R is the jet radius parameter. The jet energy (mass) response $\mathcal{R}_E$ ($\mathcal{R}_m$) is measured in bins of E_{truth} (energy of the matched truth jet) and η:

$$\mathcal{R}_E = \frac{E_{\mathrm{reco}}}{E_{\mathrm{true}}}, \qquad \mathcal{R}_m = \frac{m_{\mathrm{reco}}}{m_{\mathrm{true}}} \,. \tag{1.6}$$

The average response is determined using an iterative Gaussian fit to the core of the response distribution. The decision to calculate the response as a function of E_{truth} instead of



(a) (b)

FIGURE 1.5: The jet mass response as a function of the jet η for anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 5\%$ and $R_{\text{sub}} = 0.3$ for different jet energies (a) before and (b) after the jet mass calibration [70].

E_{reco} is motivated by the fact that for fixed $E_{\text{truth}}(E_{\text{reco}})$ bins the response distribution is (not) Gaussian. The average response is parameterized as a function of E_{reco} using a numerical inversion procedure, as detailed in Ref. [50], and the jet calibration factor is taken as the inverse of the average response. The jet mass response as a function of the jet η is depicted in Fig. 1.5 for anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ before and after the calibration procedure. After the calibration, the energy and η dependence of the jet mass response is removed and is approximately restored to one. In addition to the jet energy and mass, the jet η is also corrected with the numerical inversion technique. The standard deviation of the gaussian fit to the response distribution, σ_E (σ_m), is used to define jet energy (mass) resolution, known as JER (JMR).

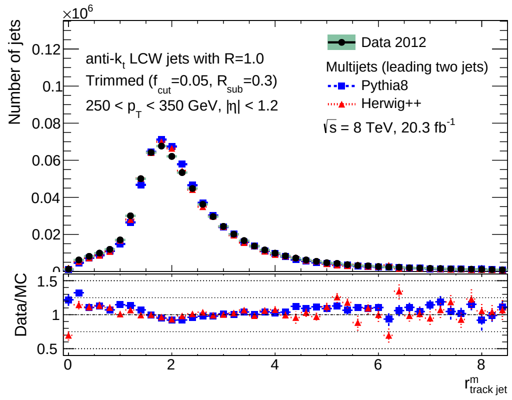
Systematic uncertainties on the JES and JMS are then derived to account for differences observed between data and simulation using well-understood reference objects:

$$\langle X^{\text{jet}} / X^{\text{ref}} \rangle_{\text{data}} / \langle X^{\text{jet}} / X^{\text{ref}} \rangle_{\text{MC}}. \tag{1.7}$$

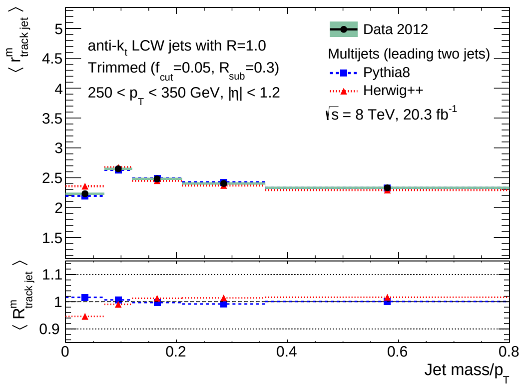
In Run-1 the JES uncertainties were derived by combining direct balance measurements in γ+jet events using photons as reference at low p_{T} with uncertainties with track-jets as reference objects in the high p_{T} regime. At the beginning of Run-2 only track-jets were used as reference objects. Uncertainties for the JMS and the scale of substructure variables (see Subsection 1.4.5) were derived only using track jets as reference objects as the γ+jet method does not provide a handle on these scales. The γ+jet studies and uncertainties are discussed in Ref. [27]. Our work focused on the track jet double-ratio method.

Track-jets were used as reference objects to derive the calibration systematic uncertainties since charged-particle tracks are well measured, independent of the calorimeter and not so susceptible to pile-up because the tracks are required to be associated to the hard-scatter vertex. Track-jets are associated to calorimeter-jets using a geometrical matching of $\Delta R < 0.75R$, where R is the radius parameter of the jet reconstruction algorithm. The double-ratio technique will be explained using the jet mass, but the procedure is the same for the jet energy and substructure variables. In this case Eq. 1.7 becomes:

$$R^{m}_{\text{track-jet}} = \frac{r^{m,\text{data}}_{\text{track-jet}}}{r^{m,\text{MC}}_{\text{track-jet}}} = \frac{\frac{m^{\text{jet}}{}_{\text{data}}}{m^{\text{track-jet}}_{\text{data}}}}{\frac{m^{\text{jet}}{}_{\text{MC}}}{m^{\text{track-jet}}_{\text{MC}}}} \tag{1.8}$$



(a)



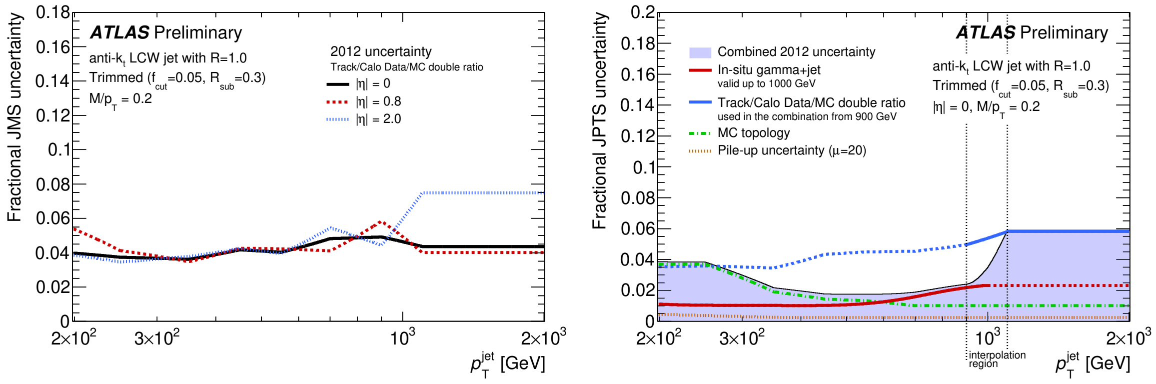
(b)

FIGURE 1.6: (a) Distribution of the calorimeter to track-jet mass ratio $r^{m}_{\text{track-jet}}$ in data and Monte Carlo simulations for anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ and (b) the mean calorimeter to track-jet mass ratio as a function of m/p_{T} [170].

Figure 1.6(a) shows the calorimeter to track-jet mass ratio, $r^{m}_{\text{track-jet}}$, distribution for the two leading anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ in PYTHIA8, HERWIG++ and the full dataset collected in 2012 corresponding to an integrated luminosity of 20.3 fb^{-1}. The peak near $r^{m}_{\text{track-jet}} \approx 2$ is caused by relatively soft particles that constitute the calorimeter jet mass. These particles are either not included in the track-jet because they result from additional pp collisions in the same bunch crossing or their tracks are bent by the magnetic field and thus result in a smaller track-jet mass. The shape of the $r^{m}_{\text{track-jet}}$ distribution is well described by both MC simulations and the mean value is also well modelled as shown in Figure 1.6(b).

Figure 1.7(a) shows the JMS uncertainty for anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ in different detector regions. The uncertainties are derived in bins of p_{T}, η and m/p_{T}. As previously mentioned, the double-ratio method as well as γ+jet events are used to derive the p_{T} scale uncertainties. The uncertainties derived from γ+jet events were however limited in p_{T} reach in the 2012 data and only statistically significant for $p_{\text{T}} < 800$ GeV. The p_{T} scale uncertainties derived with the photon-balance method are of the order of 1% at low transverse momentum and approximately 2% at high p_{T} ($p_{\text{T}} \approx 800$ GeV) compared to 4% and 6% respectively for the double-ratio track-jet method as shown in Figure 1.7(b).

A slightly different approach was used to derive JES and JMS systematic uncertainties for analyses using large-R jets that published results at the beginning of Run-2 with 3.2/fb of 13 TeV data at the end of 2015 (as is the case of the diboson resonance search in Chapter 3.2). At that point the double-ratio method systematic uncertainties were not available yet, as it takes time to derive them. The Run-1 track-jet double-ratio uncertainties shown in FIgure 1.7 were used as baseline uncertainties and additional MC uncertainties where added to take into account differences between Run-1 and Run-2. These changes include (but are not limited to) differences in the detector, the beam conditions and the reconstruction of topo-clusters, as described in Ref. [49]. These additional MC uncertainties increased the JES (JMS) by 2% (<5%) in the central region of the detector. Later analyses using the 2015 and 2016 dataset (like the diboson resonance search combination in Section 3.3 and the $Z\gamma$ measurement mentioned in Section 4.3) used in-situ JES and JMS systematic uncertainties derived using the track-jet double-ratio method described previously. The in-situ derived



(a) (b)

FIGURE 1.7: (a) Summary of the systematic uncertainties for anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 5\%$ and $R_{\text{sub}} = 0.3$ on the jet mass scale for different η ranges determined with the double-ratio track-jet method and (b) the jet energy scale derived from the track-jet double-ratio method, γ+jet events and topology differences in MC as a function of the jet p_{T}. The uncertainties are shown for $m/p_{\text{T}} = 0.2$ and were derived with data taken at $\sqrt{s} = 8$ TeV [53].

JES uncertainties are slightly smaller than the MC-based systematics used for early Run-2 analyses, while for the JMS the systematic uncertainties were reduced by about 2% overall.

Recently, at the end of Run-2, several in-situ calibration techniques where a large-R jet is balancing a photon, Z boson or another well-calibrated jet depending on the p_{T} range have been included in the large-R JES calibration chain, improving its precision [48]. The JER and a calibration of forward jets are derived using dijet balance measurements. The jet mass response is measured with two methods: using mass peaks formed by W bosons and top quarks with large transverse momenta and by comparing the jet mass measured using the energy deposited in the calorimeter with that using the momenta of charged-particle tracks. The current full calibration chain for large-R jets is presented in Figure 1.8 [48, 59]. These advanced in-situ techniques are currently used to calibrate a newest and more performant jet collection using UFO inputs and groomed using the soft-drop algorithm, which is the collection recommended for Run-3. Additional improvements in the MC-based calibration procedure are being under investigation, like the simultaneous derivation of the JES and JMS using machine learning techniques which are very promising.

1.4.5 Jet tagging

Once the jets are reconstructed and calibrated there is an extra information that might be useful when doing a physics analysis: what kind of jet is it? Is it a jet originated from a light-quark, a b-quark, a gluon, a hadronically decaying Z/W or a Higgs boson? In order to answer this question the ATLAS collaboration has developed several tagging algorithms to identify the origin of jets. In this section the main ingredients used for b-tagging and boosted boson tagging are outlined.

Boosted W/Z boson tagging

A performant boosted W/Z boson tagging requires three main ingredients:

- **A large-R jet:** At the high centre-of-mass energy of the LHC vector bosons are often produced with a transverse momentum much larger than their mass. Thus, the decay

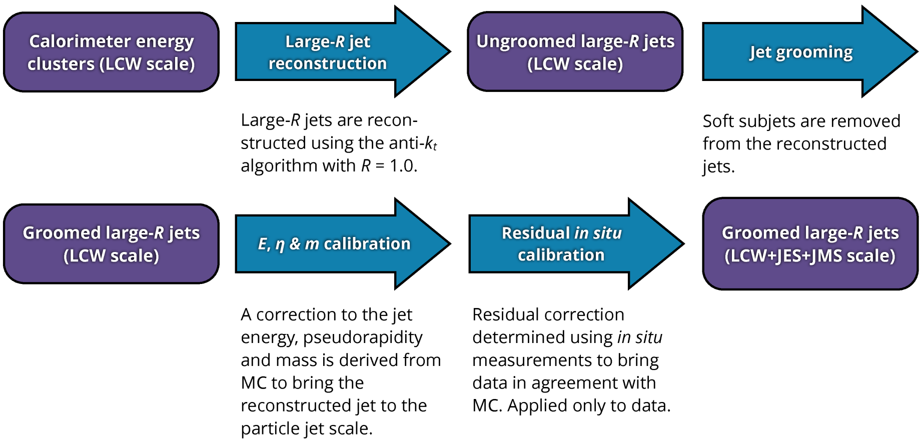


FIGURE 1.8: Stages of large-R jet energy scale calibrations at the end of Run-2. Each one is applied to the four-momentum of the jet [48].

products of the W and Z boson are boosted and collimated along the bosons momentum. The angular separation, $\Delta R_{q\bar{q}}$, of the hadronic decay products depends on the mass, $m_{W/Z}$, of the vector boson and its transverse momentum p_{T}

$$\Delta R_{q\bar{q}} \simeq \frac{1}{\sqrt{x(1-x)}}\frac{m_{W/Z}}{p_{\mathrm{T}}}, \tag{1.9}$$

where x and $(1-x)$ are the momentum fraction carried by the two quarks, respectively. For the decay of a W or Z boson, the two quarks are expected to have approximately the same momentum and thus Equation 1.9 simplifies to $\Delta R_{q\bar{q}} \simeq \frac{2\cdot m_{W/Z}}{p_{\mathrm{T}}}$. For a W boson with a p_{T} of 200 GeV, the $\Delta R_{q\bar{q}}$ between the quarks for the majority of events is approximately 0.8 whereas for a p_{T} of 500 GeV, the average angular separation is less than 0.5. Therefore, the decay products of the boosted hadronically decaying vector bosons are usually reconstructed within a large-R jet of the size of $R \approx 1.0$ to ensure that all energy deposits from the hadronic decay are captured.

- **Grooming algorithm:** a grooming algorithm is essential to remove pile-up contamination from the large-R jet. Due to the high-luminosity conditions, soft particles unrelated to the hard scattering can contaminate these jets resulting in a diminished mass resolution. A brief description of grooming algorithms can be found in Section 1.4.3.

- **Substructure variables:** Large-R jets coming from a vector boson or Higgs boson decay are characterized by two dense regions with high energy which share approximately the same p_{T} whereas for a QCD-jet (a jet from a single quark or gluon) only one dense energy region is expected assuming no hard final state radiation. *Substructure variables* are calculated from the constituents of the jet ad are designed to uncover the hard structure within jets, allowing to distinguish between large-R jets originating from the hadronic decay of bosons and those from QCD multi-jet production.

The choice of jet radius parameter size, grooming algorithm and substructure variables to use in ATLAS was extensively studied in Run-1 in ATLAS. A comparison work took place between 2014 and 2015 [40] and it is summarized in described in Section 2.3.1. This comparison study served as baseline for a boosted vector boson identification algorithm optimization for Run-2 I led in 2015 in preparation for the first $\sqrt{s} = 13$ TeV data analyses [39]. The large-R jet collection recommended and used for most of the analyses in Run-2 was

LCTopo anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$ and the studies leading to this recommendation are summarised in Section 2.3.2.

Over the past years a large number of substructure variables have been developed. These high level quantities summarize the information contained in the jet constituent's kinematic properties. Table 1.1 lists some of the variables that have been studied in ATLAS. Below only the substructure variables that will be used/mentioned in the next chapters and are particularly good at distinguish the one-prong structure of QCD-jets from the two-prong structure of boson-jets are described:

- **Jet mass:** The mass of a jet is given by the difference between the squared sums of the energy E_i and momenta p_i of the constituents, $M^2 = (\sum_i E_i)^2 + (\sum_i p_i)^2$. In the analyses described in this document three mass definitions were used for LCTopo anti-k_t $R = 1.0$ trimmed jets. The first one is the calorimeter mass, m^{calo}, is calculated using the constituent topological clusters, which are assumed to be massless. The track mass, m^{trk}, is calculated using tracks, which are assumed to have a mass of charged pions, that are ghost-associated to the jet as constituents. m^{calo} provides a good resolution for low-p_{T} jets and m^{trk} provides a good resolution for high-p_{T} jets. The two mass observables can be also combined in a third definition called combined mass, m^{comb} [48], as an inverse-resolution weighted mean using the expected jet mass resolutions determined from MC simulations. The m^{comb} algorithm provides a superior resolution, as shown Figure 1.9, and was used as the standard large-R jet mass observable in ATLAS between 2016 and the end of Run-2 when new large-R jet algorithms definitions entered at play like TCC and UFO jets.

- ***N*-subjetiness:** The N-subjettiness [215] variable, τ_{N}, describes to what degree the substructure of a jet is compatible with the hypothesis of it consisting of N or fewer subjets. τ_{N} is defined as:

$$\tau_N = \frac{\sum_i p_{\text{T},i} \min(\Delta R_{a_1,i}, ..., \Delta R_{a_N,i})}{\sum_i p_{\text{T},i} \cdot R}\,, \tag{1.10}$$

  where the sum runs over the constituents i of the given jet J, $p_{\text{T},i}$ are their transverse momenta and $\Delta R_{a_k}, i$ is the angular distance of the constituent i to the axis of subjet k. The discrimination of τ_N can be further improved with an alternative definition of the subjet axis used in Equation 1.10 [196]. The *winner-takes-all* (WTA) axis uses the hardest constituent inside the subjet k to calculate the angular separation $\Delta R_{a_k}, i$ instead of the subjet axis. The ratio of the N-subjettiness functions found with the standard subjet axes, τ_{21}, and with the WTA axes, τ_{21}^{wta}, can be used to generate the dimensionless variables that have been shown (see Figure 1.10) to be particularly useful in identifying two-body structures within jets: $\tau_{21} = \tau_2/\tau_1$ and $\tau_{21}^{\text{wta}} = \tau_2^{\text{wta}}/\tau_1^{\text{wta}}$.

- **Energy correlations:** The 1-point, 2-point and 3-point energy correlation functions for a jet are given by:

$$\text{ECF}(0,\beta) = 1, \tag{1.11a}$$
$$\text{ECF}(1,\beta) = \sum_i p_{\text{T}_i}, \tag{1.11b}$$
$$\text{ECF}(2,\beta) = \sum_{i<j\in J} p_{\text{T}_i} p_{\text{T}_j} \left(\Delta R_{ij}\right)^\beta, \tag{1.11c}$$
$$\text{ECF}(3,\beta) = \sum_{i<j<k\in J} p_{\text{T}_i} p_{\text{T}_j} p_{\text{T}_k} \left(\Delta R_{ij}\Delta R_{ik}\Delta R_{jk}\right)^\beta, \tag{1.11d}$$

TABLE 1.1: Some of substructure variables studied in the ATLAS collaboration in the last years.

Variable	Description	Reference
D_2, C_2	Energy correlation ratios	[197]
τ_{21}	N-subjettiness	[215]
R_2^{FW}	Fox-Wolfram moment	[182]
$\mathcal{P}$	Planar flow	[4]
a_3	Angularity	[15]
A	Aplanarity	[146]
$Z_{\mathrm{cut}}, \sqrt{d_{12}}$	Splitting scales	[57, 214]
$Kt\Delta R$	k_t-subjet ΔR	[142]

where the angular exponent β is a free parameter that can be optimised based on the mass of the resonance under study. Better separation of W/Z and QCD-jets was observed for $\beta = 1$ in ATLAS. Two dimensionless ratios, C_2^β [197] and D_2^β [194, 195], are defined to identify $N = 2$ dense cores of energies within the jet:

$$C_2^\beta = \frac{\mathrm{ECF}(3,\beta)\mathrm{ECF}(1,\beta)}{\mathrm{ECF}(2,\beta)^2}, \tag{1.12a}$$

$$D_2^\beta = \frac{\mathrm{ECF}(3,\beta)\mathrm{ECF}(1,\beta)^3}{\mathrm{ECF}(2,\beta)^3}. \tag{1.12b}$$

Both C_2^β and D_2^β are not infrared and collinear safe unless a mass window requirement is imposed on the jet for example around the boson mass.

- **Momentum balance:** The momentum balance variable $\sqrt{y_{12}}$ is defined as

$$\sqrt{y_{12}} = \frac{\min(p_{\mathrm{T}1}, p_{\mathrm{T}2})}{m_{12}} \Delta R_{12}\,,$$

is only meaningful for jets groomed with the split-filtering algorithm and calculated at the first step of declustering where the momentum balance (and mass-drop if specified by the algorithm) criterion is met. The momentum balance variable was used in the all-hadronic diboson search performed at $\sqrt{s} = 8$ TeV for boosted boson tagging is shown in Fig. 2.1 and described in Chapter 2.

The tagger definitions have evolved a lot in recent years. The studies you will see in Section 2.3 are based on cuts on several substructure variables. Taggers used towards the end of Run-2 combine the information from a small set of high-level summary quantities using a neural network or other machine learning (ML) methods to classify jets increasing the performance in background rejection by a factor of roughly 2-3 times [43, 87]. More recent taggers go one step further and use constituent-based information as inputs rather than high level summary substructure quantities [23]. The performance of the boson tagging algorithms in actual data is not necessarily equal to their predicted performance, for example if the jet properties are mis-modelled, so we need to apply scale factors to correct the simulated efficiency to the one measured in data. In this case the physics processes usually used for the calibration estimation are W-bosons from $t\bar{t}$ or V+jets events [86].

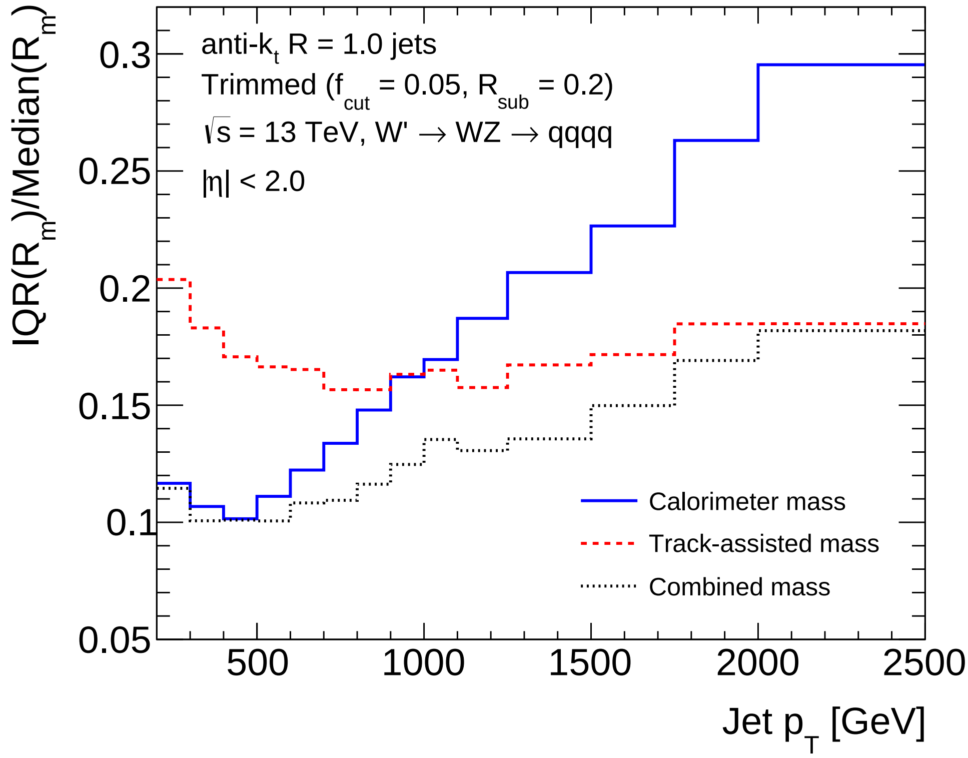
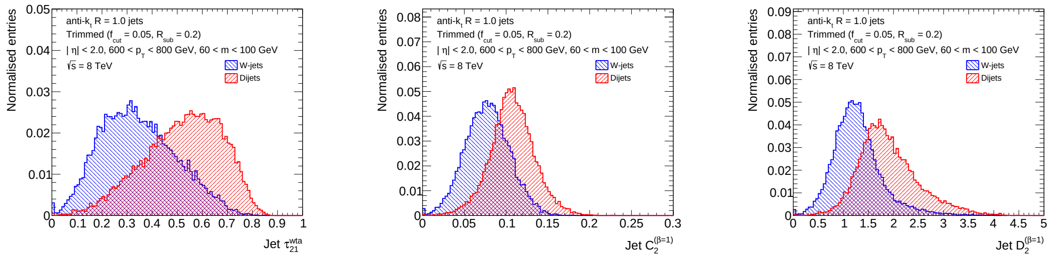


FIGURE 1.9: The fractional jet mass resolution vs the truth jet mass transverse momentum, for three different mass definitions. The resolution is calculated from LCTopo anti-k_t $R = 1.0$ trimmed jets with $f_{cut} = 0.05$ and $R_{sub} = 0.2$. The resolutions used as input to the combined mass definition are determined using a sample of jets produced via QCD dijet processes and generated with Pythia 8, where truth-jet matching is applied. The observed fractional resolutions in this figure are for jets from boosted W or Z bosons produced via $WZ \rightarrow qqqq$ processes generated using Pythia 8, with matching to particle-level W and Zs. The resolution of these boosted boson jets is expected to be different from those determined jets produced via QCD dijet processes.

b-tagging

Algorithms aiming to identify *b*-jets try to exploit the distinct features of such jets. The B hadron formed by the bottom quark has a relatively long lifetime of about 1×10^{-12} s and can travel around 3 mm before decaying. *b*-tagging algorithms exploit the fact that a certain number of tracks point to a secondary vertex instead of pointing to the reconstructed primary vertices and that the impact parameters of these tracks are large. The *b*-tagging procedure relies on the ID track reconstruction. Therefore it can be only applied to jets with $\eta < 2.5$. ID tracks are required to pass some quality criteria that depend on the different *b*-tagging algorithms. Three baseline algorithms are employed to optimize the *b*-tagging performance in the ATLAS experiment. Each of the algorithms exploits different properties of the expected signature of b-jets and thus provide complementary information. Algorithms like IP2D and IP3D [76] exploit the large impact parameters of the tracks originating from the *b*-hadron decay to determine a likelihood corresponding to the *b*-jet tag probability. The Secondary Vertex Finder algorithm (SV1) [113] attempts to reconstruct an inclusive secondary vertex. And finally JetFitter [119] aims to reconstruct the full *b*- to *c*-hadron decay chain determining a common path between the primary vertex and the *b*- and *c*-hadrons vertices (inside the *b*-jet candidate).These algorithms, first introduced during Run-1 [79],



(a) (b) (c)

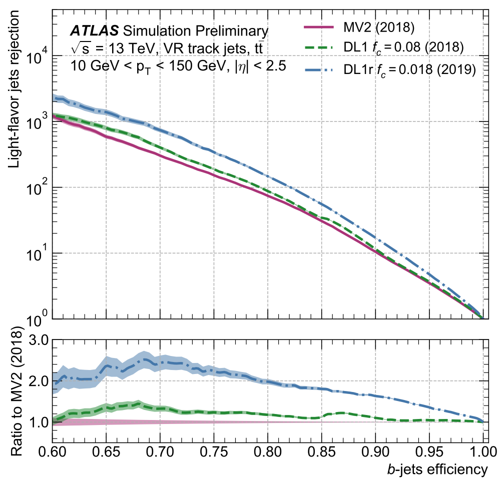
FIGURE 1.10: (a) N-subjettiness τ_{21}, (b) C_2 and (c) D_2 distribution for the leading LCTopo anti-k_t $R = 1.0$ trimmed jet with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$ in W signal and multi-jet background events with a mass window requirement [170].

benefitted from improvements and a new tuning for Run-2 [36, 76, 77].

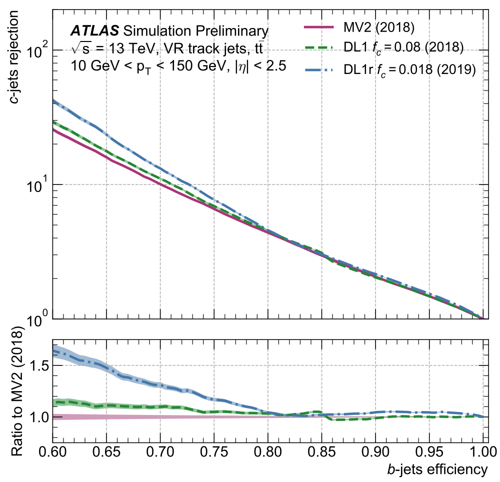
In the end, in order to maximise the b-tagging performance the outputs of these three algorithms are combined into two multivariate discriminants, namely a boosted decision tree (called MV2 [76, 77]) as well as a deep feed-forward neural network (called DL1 [76]). During the training of MV2, b-jets are set as the signal, while the background contains a mixture of 7% c-jets and 93% light-jets. Other compositions of backgrounds have been tested as well, but this particular mixture has been found to give the best separation between the different flavours. The discriminant based on this training is called MV2c10 to reflect the chosen back- ground mixture in which 'c10' represents the roughly 7% c-jets. The input variables to DL1 consist of those used for the MV2 algorithm with the addition of the JetFitter c-tagging variables dedicated properties of the secondary and tertiary vertices. One variation of the DL1 algorithm, called DL1r, additionally adds the result of the RNNIP algorithm [44], which is based on a recurrent neural network exploiting the correlation between the tracks' impact parameters.

Separate optimisations are performed for PFlow and track-jets. The MV2c10 and DL1(r) algorithm can be used at different working points (WPs). Analysers in ATLAS may choose to apply the b-tagging of jets at four different efficiencies, which are 60%, 70%, 77% and 85%. These percentages represent the fraction of b-jets which remain in simulated $t\bar{t}$ events after the corresponding cut on the MV2c10 or DL1(r) output value is applied. The lower the fraction, the fewer tagged b-jets remain, but at the same time far more c-jets, light-jets or τ-jets are rejected. Some analyses used the full tagger output information in intervals defined by the interval boundaries 0, 60, 70, 77, 85 and 100%. MV2c10 was the default algorithm in in ATLAS during Run-2 (see for example the VH analysis in Chapter 4). This recommendation changed to DL1r in 2019 for legacy physics analysis using the full Run-2 dataset. DL1 has, on average, higher light-jet rejection factors and smaller c-jet rejection factors for a given b-jet tagging efficiency with respect to the MV2 discriminants as shown in Figure 1.11.

Between October 2019 and October 2021 I was co-responsable for the Flavour Tagging calibration group. The main goal of this period was the derivation of the scale factor calibrations for newest DL1(r) recommended algorithm trained for PFlow and track jets using the full Run-2 dataset for physics analyses [13, 60, 71, 78]. The nominal MC samples may have a different performance from MC samples used in physics analyses, and they can not fully describe data.Therefore, the MC-to-MC scale factors and the data-to-MC scale factors are needed. b-jets tagging efficiency was measured using $t\bar{t}$ events where both top quarks decay leptonically, $c-$tagging efficiency from $W \to cs$ decay in $t\bar{t}$ events and the light-jet



(a)



(b)

FIGURE 1.11: Comparison of ROC curves for 2018 recommended versions of MV2 and DL1, and 2019 DL1r optimization. The x-axis corresponds to the b-jet efficiency, while the y-axis corresponds to (a) light-flavor jets rejection ($1/\epsilon_u$) and (b) c-jets rejection ($1/\epsilon_c$). The performance is evaluated on track-jets with $10 < p_{\mathrm{T}} < 150$ GeV and $|\eta| < 2.5$, on a simulated dataset of $t\bar{t}$ events. The shaded bands represent the statistical uncertainty [44].

mistag efficiency from Z+jets events with enriched light-flavour jets. Several other alternative calibration analyses (high p_{T} calibrations using dijet events, charm mis-tag rate calibration based on $W + c$ events, $b - jet$ semileptonic $t\bar{t}$ calibration in boosted regime using a tag and probe method, τ calibration, boosted $H \to b\bar{b}$) also progressed during this time. I also overview the migration of the flavour tagging calibration code to release 22, in preparation to the start of Run-3.

Boosted $H \to b\bar{b}$ tagging

After working in boosted W/Z tagging and with the idea of expanding the hadronically decaying diboson searches to topologies including Higgs($\to b\bar{b}$) bosons, I was co-responsable of the (at the time) newly created $X \to bb$ performance subgroup between mid-2015 and mid-2016. During that time we provided the first recommendations for tagging boosted Higgs bosons in the first Run-2 data in ATLAS. The recommendations were based directly on a large body of work carried out in both real and simulated data from Run-1 [38, 81], and throughly re-evaluated and extended in simulated data at $\sqrt{s} = 13$ TeV and documented in Ref [35]. Later on I focused on the first full comparison study, similar to the one performed for W/Z jet tagging, that lead to the tagger recommendation used for most of the ATLAS physics analysis in Run-2: the two b-tagged track-jet tagger, to be discussed below. This work was documented in a conference note [17] and a paper published in 2019 which I co-edited [41].

In ATLAS there are currently two main techniques to identify boosted Higgs-jets:

- **Two b-tagged track-jet tagger:** In its basic form this tagger consist in reconstructing the

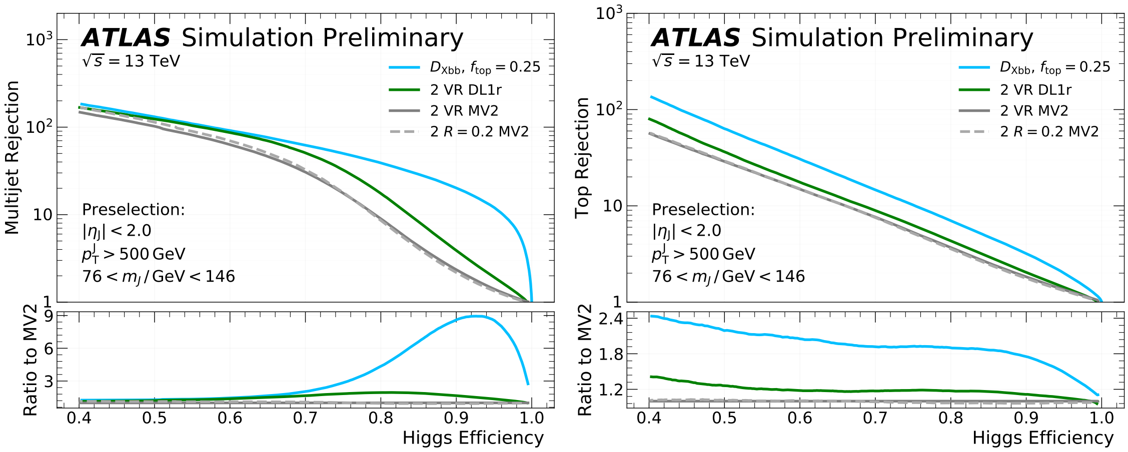
Higgs boson candidate as a large-R jet and applying b-tagging requirement to track-jets ghost-associated [82, 135, 138] to the large-R jet in order to select candidates corresponding to $H \to b\bar{b}$ decays. The initial recommendations used track-jets built with the anti-k_t algorithm with a radius parameter of $R = 0.2$ [35, 38]. $R = 0.2$ jets have a limited performance at high p_T and therefore at the beginning of Run-2 recommendations moved to use a variable radius (VR) p_T-dependent parameter, from tracks reconstructed in the inner detector with $p_T > 0.5$ GeV and $|\eta| < 2.5$ [122, 193]. VR track-jets have an effective jet radius R_{eff} proportional to the inverse of the jet p_T in the jet finding procedure: $R_{\text{eff}}(p_T) = \rho / p_T$, where the ρ-parameter is set to 30 GeV. There are two additional parameters, R_{min} and R_{max}, used to set the minimum and maximum cutoffs on the jet radius, and these are set to 0.02 and 0.4, respectively. Standard b-tagging algorithms with a dedicated training for track-jets are used, being the most popular algorithms: MV2c10 and DL1r. As described in the previous section the performance and calibrations are based on isolated track jets and jet flavours. More complicated versions of this tagger were also studied by requiring the b-tagged large-R jet mass to be around the SM Higgs boson mass of 125 GeVand a requirement on other large-R jet substructure variables [41].

- **Neural networks (NN) tagger:** this description is taken from Ref. [30]. A neural network (NN) is trained using the large-R jet p_T and η, along with the discriminants from the DL1r algorithm of up to three associated VR track-jets [42]. The large-R jet collection used was anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$ and topological clusters with noise suppression as inputs. The $X \to b\bar{b}$ tagger provides three outputs for each large-R jet, corresponding to the probabilities assessed of being a Higgs boson (p_{Higgs}), multijet (p_{multijet}) and top-quark (p_{top}) jet. A single discriminant, roughly corresponding to a log-likelihood ratio, is defined:

$$D_{\text{Xbb}} = \ln \frac{p_{\text{Higgs}}}{f_{\text{top}} \cdot p_{\text{top}} + (1 - f_{\text{top}}) \cdot p_{\text{multijet}}}, \tag{1.13}$$

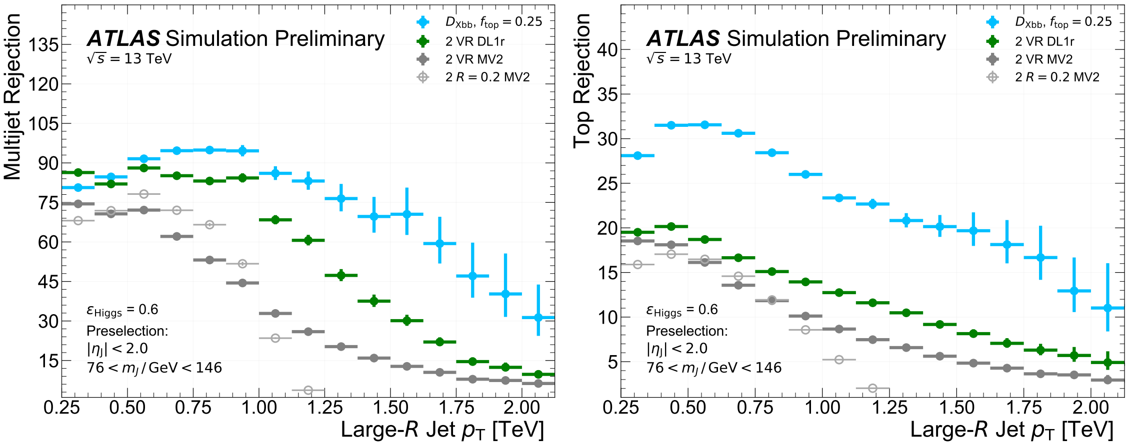
  where f_{top} determines the fraction of the top-quark background. The tagging of a large-R jet as originating from an $X \to b\bar{b}$ process is therefore achieved by imposing a selection on the D_{Xbb} discriminant. Three WPs corresponding to 50%, 60% and 70% $X \to b\bar{b}$ tagger signal efficiencies with $f_{\text{top}} = 0.25$ are recommended for Run-2 analyses. Figures 1.12 compares the ROC efficiency curves for the NN tagger and the two b-tagged track-jet tagger using $R = 0.2$ and VR track-jets. Figure 1.13 shows the multijet (left) and top (right) jet rejection as a function of large-R jet p_T with a constant 60% Higgs tagging efficiency.

During my time as co-responsable of the $X \to b\bar{b}$ performance subgroup we also started exploring possible in-situ calibrations for the Higgs-tagger, instead of using the calibrations based on isolated track jets and jet flavours. Our aim was to develop new calibration methods that took into account the full correlations between the tagger variables. In 2016 I oversaw the first b-tagging performance and jet substructure studies using a high p_T $g \to b\bar{b}$ rich sample of large-R jets [114]. Around the same time I started co-supervising Miles Wu, (at the time) PhD student at University of Chicago (2016-2017). We explored the possibility of using $Z \to b\bar{b}$ events for the tagger calibration. The calibration based on $Z(\to b\bar{b})\gamma$ events was first considered by ATLAS in our paper Ref. [41] and a first jet mass measurement using the same events was documented in Ref. [61]. Later on, during my tenure as co-responsable of the Flavour Tagging calibration subgroup, I oversaw the first set of calibration for the $H \to b\bar{b}$ tagger ever, recently object of a public ATLAS note [30]. The calibration of the NN tagger is performed according to a process-based approach, as opposed to the flavour-based



(a) (b)

FIGURE 1.12: (a) Multijet and (b) top jet rejection as a function of the $H \rightarrow b\bar{b}$ tagging efficiency, for large-R (R=1.0) jet $p_{\mathrm{T}} > 500$ GeV. Performance of the DXbb algorithm is compared to DL1r and to two variants of MV2, one evaluated on variable-radius (VR) jets, the other on fixed-radius (FR, $R = 0.2$) jets [42].



(a) (b)

FIGURE 1.13: (a) Multijet and (b) top jet rejection as a function of large-R jet p_{T} with a constant 60% Higgs tagging efficiency. Error bars correspond to the statistical uncertainty. Fixed-radius track-jets with $R = 0.2$ begin to lose discriminating power when $p_{\mathrm{T}} > 800$ GeV, since subjets begin to merge and resolving two independent b-jets becomes difficult. Variable-radius track-jets, by contrast, can resolve multiple subjets in boosted objects with $p_{\mathrm{T}} > 1$ TeV. [42].

approach typically pursued in the calibration of track-jets. Thus, the signal efficiency calibration is performed with $Z(\to b\bar{b})\gamma$ and $Z(\to b\bar{b})$+jets events at high Z-boson transverse momentum and the mistagged efficiency is studied for the top and multijet ($g \to b\bar{b}$) processes. I worked actively in the $Z(\to b\bar{b})$+jets calibration, together with Yajun He, a PhD student at LPNHE who I co-supervise since 2019. More details in Chapter 4.

1.5 Chapter summary and outlook

A better understanding of our detectors is essential to achieve the precision measurements and searches that are part of the LHC research program and this is only achievable through detector performance work. If we follow the evolution of the last years then we can expect huge performance improvements in the years to come, in particular in the jets and QCD sector, thanks to increased statistics, exploration of new physics processed for objects calibration and of more advanced techniques in the machine learning-world. These improvements will only be achieved thanks to the continuous commitment and expertise developed by our scientists in detector performance activities. Visibility and recognition for this work are fundamental.

Chapter 2

Searching for diboson resonances in Run-1

The substantial dataset of the LHC proton-proton (pp) collisions at $\sqrt{s} = 8$ TeV collected by the ATLAS experiment towards the end of Run-1 provided a distinct opportunity to search for new heavy resonances at the TeV mass scale. This chapter presents the results of a search for narrow diboson resonances (WW, WZ and ZZ) decaying to fully hadronic final states using a dataset of 20.3/fb at $\sqrt{s} = 8$ TeV. Section 2.1 is largely based on Refs. [99] and [170]. The analysis suffers from the enormous background of SM QCD dijet production. Further contributions from SM processes, all of which are much smaller in magnitude at high p_{T}, are W/Z+jets, diboson, top-quark pair and single-top production. The dominant QCD background was significantly suppressed by reconstructing the decay products of each vector boson as one single large-R jet and then applying the grooming and tagging techniques based on jet mass and substructure properties. The signature of the decay of a new heavy particle is thus a narrow resonant structure on top of the smoothly falling invariant dijet mass spectrum of the background processes. The largest deviation from a smoothly falling background in the observed dijet invariant mass distribution occurs around 2 TeV in the WZ channel, with a global significance of 2.5 standard deviations. Excesses like have been observed in other diboson analyses by CMS too using the 2012 dataset [161]. These results attracted interest in the community at that point of time and set high expectations on the Run-2 analysis. This is why I spent a significant amount of time in the optimisation of the jet boson tagging optimisation in 2014 as described in Section 2.3. Section 2.3 is largely based on Refs. [40] and [39].

I joined the Run-1 search for hadronically decaying diboson resonances analysis at the end of 2013 during my postdoc at University of Geneva. I was involved in this analysis as one of the main analysers and as the co-supervisor of (at a time) a PhD student in University of Geneva, Chris Delitzsch. Among the contributions to the analysis one could mention the calibration of the large-R jets, its corresponding uncertainties and detailed investigations of the excess of events in the 2012 data, including but not limited to pile-up dependency, data-taking stability, calorimeter energy distributions and calorimeter punch-through studies.

2.1 Hadronically decaying diboson resonance search at 8 TeV

2.1.1 Theoretical and experimental context

The SM of particle physics predicts the experimental observations over a broad range with remarkable success. Nonetheless, several experimental observations cannot be explained by the SM and open questions remain:

- Why is the Higgs boson so light? Related with the so-called naturalness problem.
- What is the origin of the matter-antimatter asymmetry?

- Why 3 fermion families? Why do neutral leptons, charged leptons and quarks behave differently?
- What is the origin of neutrino's mass and oscillations?
- What is the composition of dark matter?
- What is the cause of the Universe's accelerated expansion?

This require an extension of the theoretical framework. Several extensions have been proposed in the last years but so far none of them have managed to cover and/or solve all the limitations of the SM. Those theories are known as Beyond-the-SM (BSM) theories and their verification/refutation is one of the goals of the LHC physics program. Dibosons are standard topologies when looking for new physics, since they are linked to the electroweak symmetry breaking mechanism. Diboson resonances are predicted by different BSM theories such as technicolour [7, 144, 174], warped extra dimensions [209, 210], and Grand Unified Theories [183, 184]. To assess the sensitivity of the search presented in this chapter to BSM effects, to optimise the event selection, and for comparison with data, two specific benchmark models were used: an extended gauge model (EGM) $W' \rightarrow WZ$ where the spin-1 W' gauge boson has a modified coupling to the SM W and Z bosons [5], and a spin-2 graviton, $G_{RS} \rightarrow WW$ or ZZ, a Kaluza–Klein mode of the bulk Randall-Sundrum (RS) graviton [209].

Up to 2012 previous searches were mostly performed in the fully leptonic or semileptonic decay modes because leptons are well-understood objects and provide clean signatures. However the development of boosted boson tagging techniques, introduced in Section 1.4.5, opened the possibility to exploit the fully hadronic diboson modes as well. The fully hadronic mode has a higher branching fraction than leptonic and semileptonic decay modes, and is therefore used to extend the reach of the search to the highest possible resonance masses.

The panorama of diboson resonances searches at LHC at the time of this analysis is summarised in the following. The ATLAS Collaboration had explored the $WW/WZ \rightarrow \ell \nu jj$ final state [107] using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV. For the extended gauge model W' boson, masses below 950 GeV could be excluded at 95% confidence level, exceeding the limits from the Tevatron collider at Fermilab [129, 207]. For the bulk RS graviton masses below 710 GeV were excluded. Regarding $\sqrt{s} = 8$ TeV analyses, using the $\ell\ell q\bar{b}$ final state, the ATLAS collaboration has excluded at 95% CL a bulk $G_{RS} \rightarrow ZZ$ with mass below 740 GeV [105]. For narrow resonances decaying exclusively to WZ or WW, the sensitivity of the ATLAS search in the $\ell \nu q\bar{q}$ channel [102] is comparable to that of the search presented here in the mass range from 1.3 to 2.5 TeV. That search has also excluded a bulk $G_{RS} \rightarrow WW$ with mass below 760 GeV.

The CMS Collaboration searched for VZ resonances using the $\sqrt{s} = 7$ TeV dataset [157] where the Z boson decays into lepton pairs and the V into two overlapping jets and for the first time as well as for the fully hadronic decay of VV resonances [158]. For the semileptonic search, W' masses between 700 and 940 GeV as well as graviton masses between 750 and 880 GeV were excluded assuming $k/\bar{M}_{\rm Pl} = 0.05$ (the strength of the coupling depends on $k/\bar{M}_{\rm Pl}$ where k corresponds to the curvature of the warped extra dimension and $\bar{M}_{\rm Pl}$ is the four-dimensional Planck scale). The fully hadronic search however was only sensitive to set upper limits on the production cross-section times branching ratio. Using the $\sqrt{s} = 8$ TeV dataset CMS performed a search for diboson resonances with the fully hadronic final state [160] of comparable sensitivity to the one presented in this Chapter. In this search, the EGM $W' \rightarrow WZ$ with masses below 1.7 TeV and G_{RS} of the original RS model decaying to WW with masses below 1.2 TeV are excluded at 95% confidence level (CL). The CMS collaboration has also published upper limits on the production of generic diboson resonances using semileptonic final states [159, 161]. The largest deviation from the background model was found in the $G_{RS} \to WW$ selection at a mass of 1.9 TeV with a local significance of 1.3.

2.1.2 Monte Carlo simulation and data

This search was performed using the dataset collected in 2012 from 8 TeV LHC pp collisions using a single-jet trigger with a nominal p_{T} threshold of 360 GeV. The integrated luminosity of this dataset after requiring good beam and detector conditions is 20.3/fb, with a relative uncertainty of $\pm$2.8% [46].

$W' \to WZ \to qqqq$ signal events, predicted in an extended gauge model (EGM) are generated with PYTHIA8 [213] using the MSTW2008 [202] PDFs and the AU2 [115] underlying event tune. The Randall–Sundrum graviton signal samples are generated in Run-1 with CALCHEP [126] using the CTEQ6L1 [205] parton distribution functions and PYTHIA8 for the simulation of the hadronisation and the parton shower. The RS model is characterised by $k/\bar{M}_{\mathrm{Pl}} = 1.0$ and the graviton can either decay into WW or ZZ whereupon only the hadronic vector boson decay is simulated. The signal samples are generated with different W'/G_{RS} masses, covering the range $1.3 < m_{W'} < 3.0$ TeV in 100 GeV intervals.

The QCD dijet background events are produced with PYTHIA8 and the CT10 parton distribution functions. The final-state particles produced by the generators are propagated through a detailed detector simulation [117] based on GEANT4 [2]. The expected contribution from additional pp interactions is accounted for by overlaying minimum-bias events on the events of interest. These additional events are generated with PYTHIA8 and the A2 tune. The distribution of the average number of pp interactions per bunch crossing in MC simulation was adjusted to the distribution in data, with an average of 20.7. Simulated events are then reconstructed with the same algorithms run on collision data.

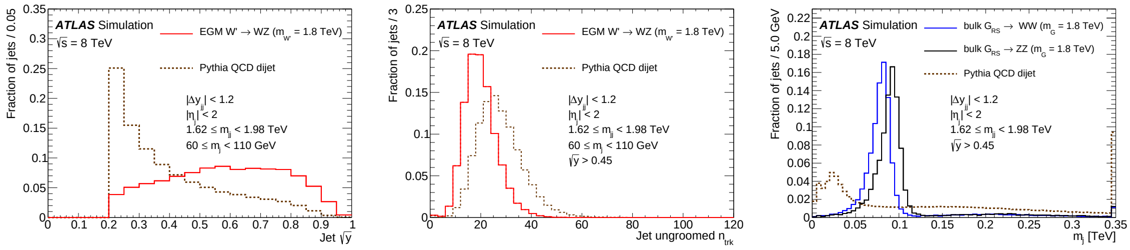
2.1.3 Object and event selection

Jets are reconstructed with the C/A algorithm using a radius parameter of $R = 1.2$ from calorimer-only information. To remove contamination from pile-up and the underlying event, the split-filtering algorithm as described in Section 1.4.3 is used with the parameters $\sqrt{y_{\mathrm{cut}}} = 0.2$ and $\mu_{\max} = 1.0$. In the filtering stage, the jet constituents at the last step of de-clustering procedure are reclustered using a fixed subjet radius size of $R_{\mathrm{sub}} = 0.3$ and only the three highest-p_{T} subjets are kept. The energy and mass of the large-R jets are then calibrated following the procedure described in Section 1.4.4.

Events are required to pass the single high-p_{T} jet trigger, introduced in the previous section, and to have a primary vertex with at least two tracks with $p_{\mathrm{T}} > 500$ MeV. Events are removed if any muon (loose quality, $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$) or electron (medium quality, $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.47$ excluding the transition region between the barrel and endcaps in the LAr calorimeter) is present or if $E_{\mathrm{T}}^{\mathrm{miss}} \geq 350$ GeV, ensuring that this analysis has no events in common with other diboson search analyses [102, 105]. Two large-R groomed jets with $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.0$ are required. The η cut ensures sufficient overlap with the ID, since associated charged-particle tracks are used in the boson tagging requirements and for estimating systematic uncertainties. To select only events in the region where the trigger is 99.9% efficient, the leading-p_{T} jet is required to fulfil $p_{\mathrm{T}} > 540$ GeV. To avoid distortions of the invariant mass spectrum of the two leading jets (m_{JJ}) due to the p_{T} requirements, only events with $m_{JJ} > 1050$ GeV are selected.

Then two topology cuts, optimised for the signal samples with masses 1.2, 2.0 and 3.0 TeV, are applied:

- The rapidity difference between the two leading jets is required to be $|\Delta y_{12}| = |y_2 - y_1| < 1.2$. This rapidity difference is smaller for s-channel processes such as the W'



(a) (b) (c)

FIGURE 2.1: Comparison of the boson-tagging variables used in the 8 TeV diboson analysis for signal jets and QCD dijet background: (a) subjet momentum balance $\sqrt{y_{12}}$, (b) number of tracks matched to the jet n_{trk} and (c) jet mass. The signal and background distributions are normalised to unit area, and the last bin of each histogram includes the fraction of events falling outside of the displayed range [99].

and G_{RS} signal models than for the t-channel processes dominating the QCD dijet background.

- The p_{T} asymmetry between the two leading jets is required to be $A = \frac{p_{\mathrm{T},1} - p_{\mathrm{T},2}}{p_{\mathrm{T},1} + p_{\mathrm{T},2}}$, where $p_{\mathrm{T},1} < 0.15$. It is used to reject events where one of the jets is poorly measured or does not come from the primary pp collision, since for the two-body decay of a resonance, the decay products are expected to be balanced in momentum.

The two leading p_{T} jets each must satisfy the following boson tagging requirements:

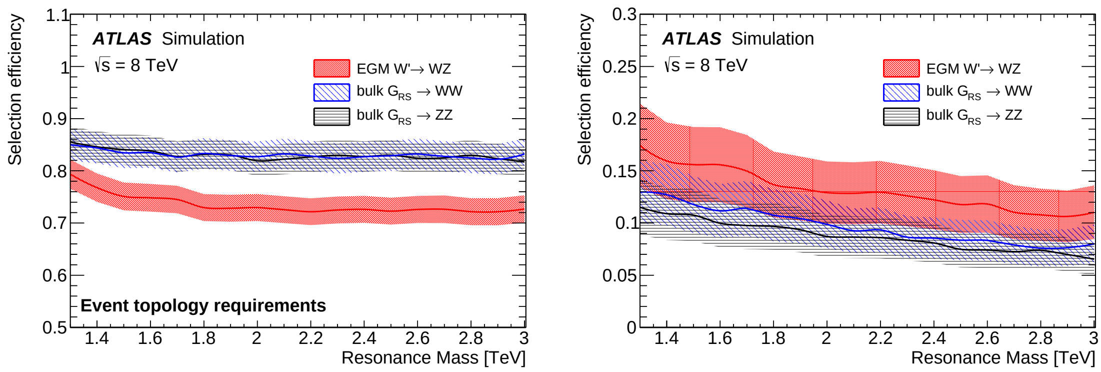
- A more stringent momentum balance criterion is applied, $\sqrt{y_{12}} > 0.45$, since jets in QCD dijet events that survive grooming tend to have unbalanced subjet momenta characteristic of soft gluon radiation.

- The number of charged-particle tracks associated with the original, ungroomed jet is required to be small, $n_{\mathrm{trk}} < 30$. Due to the larger colour factor of gluons ($C_A = 3$) compared to quarks ($C_F = 4/3$), a larger number of particles is produced in the fragmentation of a gluon.

- A selection window is applied to the invariant mass of the filtered jet, m_J, since this quantity is expected to be small for jets in QCD dijet events and to reflect the boson mass for jets from hadronic boson decays. Only jets within a 26 GeV wide mass window, centred around the peak of the W-jet and Z-jet mass distribution in MC simulation, are selected. The peaks of the W-jet and Z-jet mass distributions correspond to 82.4 GeV and 91.8 GeV, respectively.

The $\sqrt{y_{12}}$, n_{trk} and jet mass distribution for signal and background events is shown in Figure 2.1. Finally, the invariant mass calculated from the two leading jets must exceed 1.05 TeV. Figure 2.2 shows the selection efficiency of the event topology requirements for signal events with resonance mass within 10% of the nominal signal mass for the $W' \to WZ$, bulk $G_{RS} \to WW$ and bulk $G_{RS} \to ZZ$ benchmark models, with statistical and systematic uncertainties indicated by the width of the bands in the figure.

The efficiency of the n_{trk} requirement was corrected by a scale factor derived from data since the variable is not well modeled in MC. A scale factor of 0.90 ± 0.08 was derived from the ratio of the selection efficiency of this cut in a data control region enriched with W/Z+jets events, where a high-p_{T} W or Z boson decays hadronically, to the selection efficiency of this cut in simulation.



(a) (b)

FIGURE 2.2: Signal selection efficiencies as a function of the resonance mass for the $W' \to WZ$ and the bulk graviton models $G_{\mathrm{RS}} \to WW, ZZ$ after requiring (a) the event topology selection only and (b) after the WZ, WW and ZZ boson tagging selections are also applied [99].

2.1.4 Background estimation

The background in the analysis is not estimated from MC simulation but rather from the data using a parametric function to avoid statistical limitations and mismodelling of the background MC samples especially in the high mass region. The background for this search is largely dominated by QCD dijet events and the contributions from other SM processes such as diboson production, W/Z+jets and $t\bar{t}$ events are either small or negligible and not expected to distort the shape of the background distribution. The shape of the background is determined by performing a maximum-likelihood fit to the observed invariant mass m_{JJ} spectrum using the following empirically characterised by the function:

$$\frac{\mathrm{d}n}{\mathrm{d}x} = p_1(1-x)^{p_2+\xi p_3}x^{p_3}\ , \quad x = m_{JJ}/\sqrt{s}\,, \tag{2.1}$$

where p_1 is a normalisation factor, p_2 and p_3 are dimensionless shape parameters and ξ is a dimensionless constant chosen after fitting to minimise the correlations between p_2 and p_3. To estimate the background, the fit is performed in the range $1.05 < m_{JJ} < 3.55$ TeV, where the lower limit is dictated by the point where the trigger is fully efficient for tagged jets and the upper limit is set to be in a region where the data and the background estimated by the fit are well below one event per bin for the tagged distributions.

The background functional form was tested for compatibility to describe distributions with similar characteristics: the invariant dijet mass in MC generated QCD dijet events for different MC generators (PYTHIA8, HERWIG++) before and after boosted boson tagging requirements, also in collision data events in a phase-space close to the signal region like at different boosted boson tagging stages, and in otherwise tagged events where both the leading and subleading jet have masses falling below the boson-tagging mass windows. The dijet mass distribution of the simulated background and of each of these background-dominated selections are well-described by the functional form in eq. 2.1. Further studies were performed to determine if the background determination could be significantly improved by adding additional degrees of freedom to the functional from. The goodness of fit for the different functional forms was tested as a function of the collected data sample size with a log-likelihood ratio statistics employing Wilk's theorem [218]. These studies showed that no additional degrees of freedom were necessary for the luminosity studied in this version of the analysis.

2.1.5 Results

Background fit to data

Figure 2.3 shows the fitted dijet mass distribution of events passing the WW, ZZ, WZ and also the combined three tagging selections using the fitting procedure described in Section 2.1.4. The size of the shaded band reflects the uncertainties on the fit parameters propagated to show the uncertainty on the expectation from the fit. These uncertainties were probed to cover differences due to the use of alternative families of parametrisations, and by considering signal plus background fits of the chosen function to simulations of the dominant background as well as sidebands and control regions of data in which a signal contribution is expected to be negligible. The total number events observed was 425, 333 and 604 in the WW, ZZ and WZ signal region, respectively. Due to the large-R jet mass resolution and therefore the size of the chosen mass window around the W boson and Z boson mass, the three signal regions are not orthogonal and thus about 20% of the selected events are shared among all three signal regions. The fraction of events selected by both the WZ and WW selection are 49% whereas 43% of the events are common to the WZ and ZZ selection.

The m_{JJ} distribution is well described by the background model except for a few bins around 2 TeV. The compatibility of the data with the background model was tested in terms of the local p_0 value, shown in Figure 2.4 for the three signal regions. The largest discrepancy was observed in the WZ signal region with a local significance of 3.4σ and a global significance of 2.5σ (taking into account that the three signal regions are not independent).

Additional cross-checks

Several cross-checks were performed to ensure that the excess of events observed around $m_{JJ} = 2$ TeV was not caused by a bias in the reconstruction techniques or event selection used. Some of these cross-checks will be summarised here. The events were not accumulated in a particular region in the detector, as shown in Figure 2.5. Furthermore, no dependence on the number of primary vertices, the average number of interactions or the data taking period was observed. Event displays for each one of the events with m_{JJ} close to 2 TeV were produced and analysed, no obvious biased was found. Figure 2.6 show two of those events.

Another test was to study the m_{JJ} spectra when varying the different signal regions by removing one of the boson tagging requirements as shown in Figure 2.7 for the WZ signal region case. Only the combination of the mass window and the momentum balance cut results in the excess of events around 2 TeV. On the other hand, when only one boson tagging criterion is applied at a time, each of the m_{JJ} spectra is a smoothly falling distribution (see also Figure 2.7).

The choice of jet reconstruction and grooming algorithm was considered as one possible source of distortion in the invariant mass spectrum. Unfortunately at the time of the analysis no other algorithms were studied in depth, optimized and calibrated. After the Run-1 analysis finished, a full optimisation study for boosted boson tagging was performed (more details in next Section 2.3), yielding as recommended algorithm the LCTopo anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{sub} = 0.2$. Dedicated energy and mass calibrations were derived for this algorithm at $\sqrt{s} = 8$ TeV to be able to re-analyse the 2012 data. Figure 2.8 shows the m_{JJ} distributions for the WW, ZZ and WZ regions using the Run-2 algorithm. Two things can be noted, the background rejection is higher and a slight excess of events can be observed for the WZ and ZZ selection but with a much smaller significance.

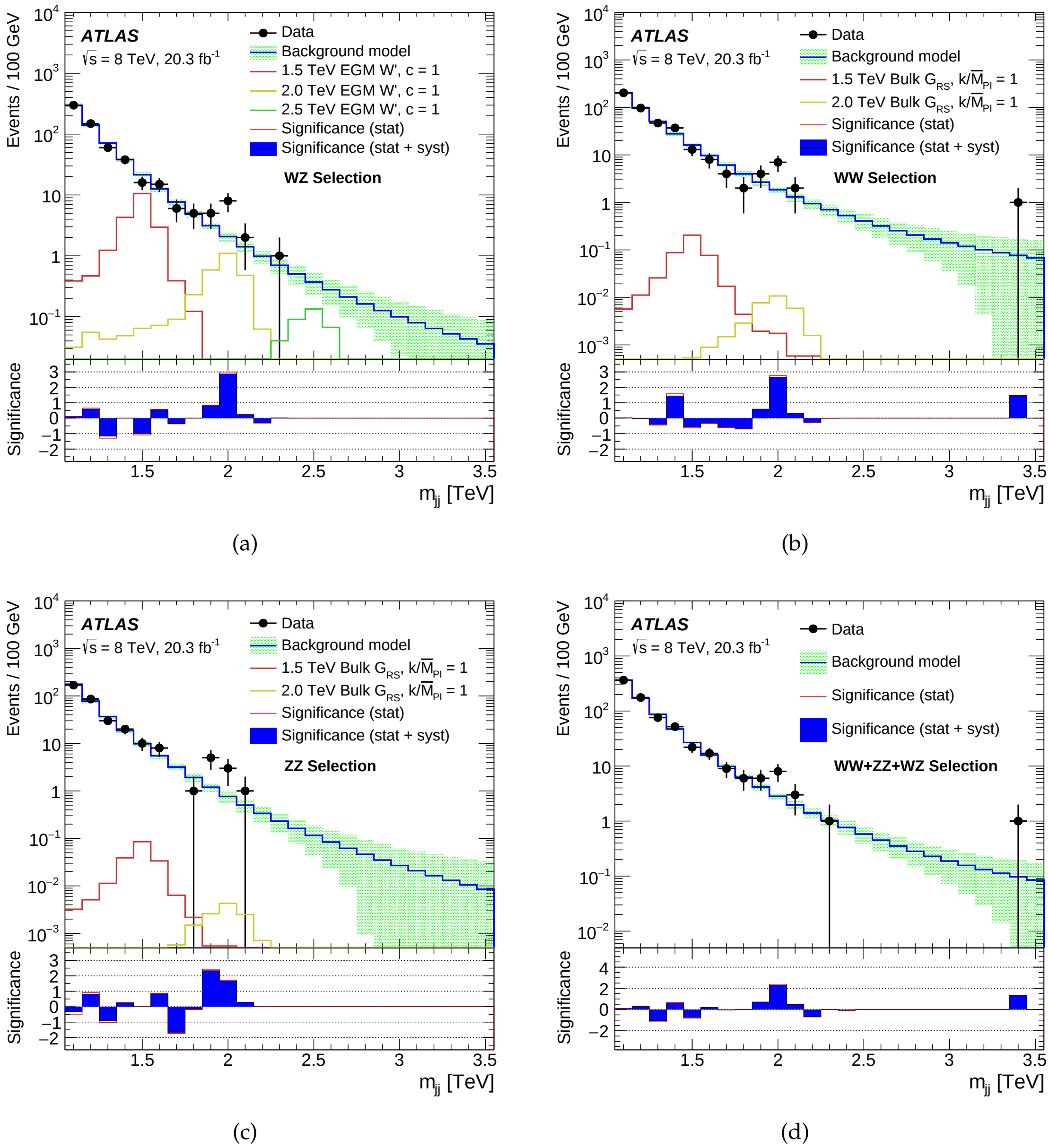


FIGURE 2.3: Background-only fit to the invariant dijet mass (m_{JJ}) distribution in $\sqrt{s} = 8$ TeV data for boson-tagged jets in the three different signal regions: (a) WZ, (b) WW, (c) ZZ and (d) for events passing any of the three tagging selections. The significance shown in the inset for each bin is the difference between the data and the fit in units of the uncertainty on this difference. The significance with respect to the maximum-likelihood expectation is displayed in red, and the significance when taking the uncertainties on the fit parameters into account is shown in blue. The spectra in the three signal regions are compared to the signals expected for an W' with $m_{W'} = 1.5,\ 2.0$ or 2.5 TeV in the WZ signal region and a RS graviton with $m_{G_{RS}} = 1.5$ or 2.0 TeV in the WW and ZZ signal region [99].

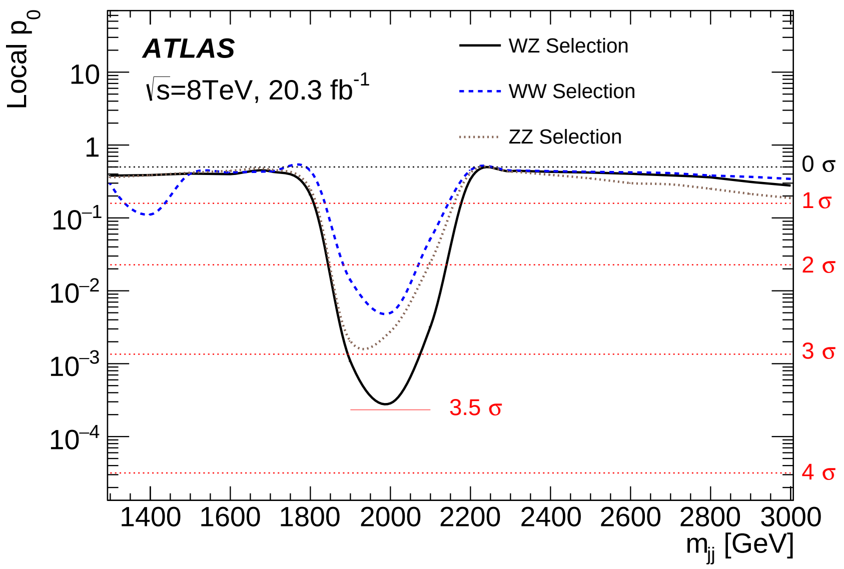


FIGURE 2.4: The local p_0 value and local significance for the WZ, WW and ZZ signal regions. Since the WZ, WW and ZZ selections are not orthogonal, the observed p0 are not independent between the three selections [99].

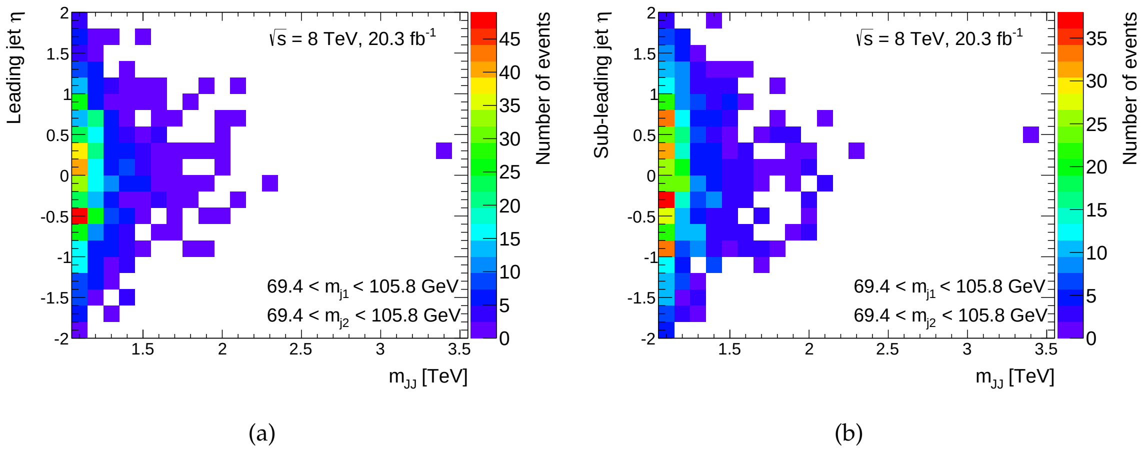
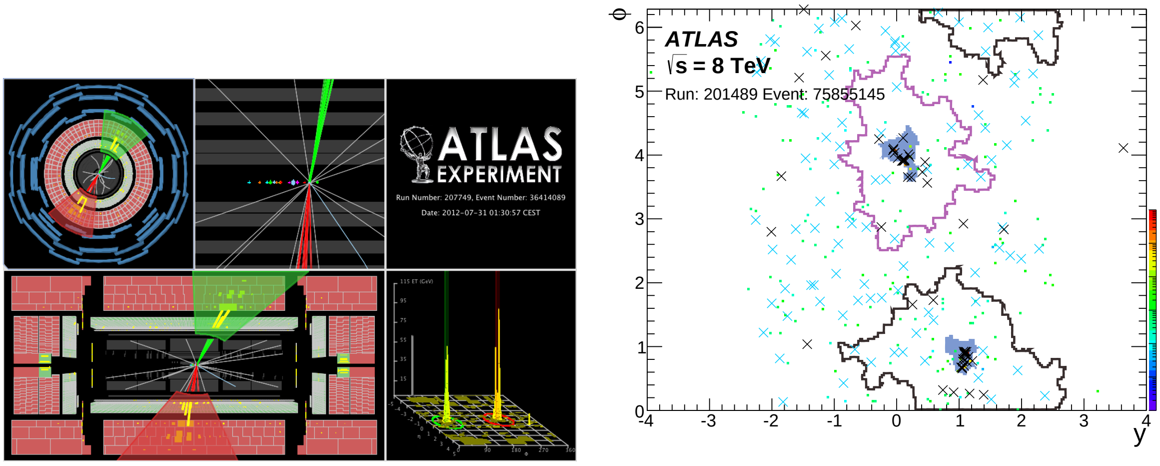
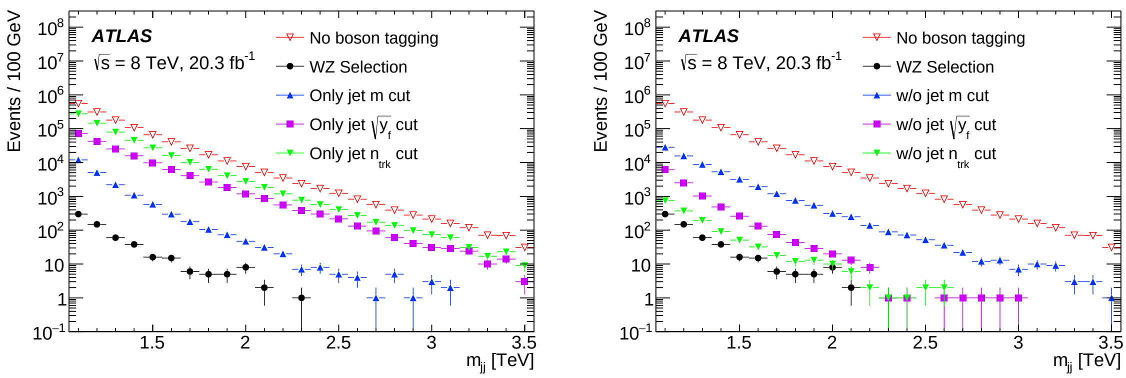


FIGURE 2.5: η distribution of the (a) leading and (b) subleading jet as a function of the invariant dijet mass in a region in $\sqrt{s} = 8$ TeV data covering the three signal regions [170].



(a) (b)

FIGURE 2.6: (a) Event Display of Event 36414089 in Run 207749, which is selected by the WZ selection. The dijet invariant mass m_{JJ} is 2.26 TeV, the leading jet $p_{\mathrm{T}} = 1.1$ TeV and $m_j = 104$ GeV, and the subleading jet has $p_{\mathrm{T}} = 1.1$ TeV and $m_j = 71$ GeV. (b) Jet display for Event 75855145 in Run 201489. The colored boxes represent the topological clusters, the blue crosses represent tracks not associated with the event primary vertex, the black crosses represent tracks associated with the event primary vertex, and the gray areas are the reclustered jets kept after the filtering algorithm. $p_{T,1} = 828.4$ GeV; $\eta_1 = 1.1$; $\phi_1 = 0.82$; $m_1 = 97.9$; $n_{\mathrm{trk},1} = 23$; $p_{T,2} = 818.2$ GeV; $\eta_2 = 0.09$; $\phi_2 = -2.3$; $m_2 = 91.8$; $n_{\mathrm{trk},2} = 28$; $m_{JJ} = 1873.2$ GeV.



(a) (b)

FIGURE 2.7: Observed m_{JJ} spectra when varying the WZ boson tagging selection by removing requirements. In (a), the spectrum with no tagging and the spectrum with the full WZ selection are compared to spectra with just one boson tagging requirement applied to each jet. In (b), the effect of applying all tagging requirements except one is displayed [170].

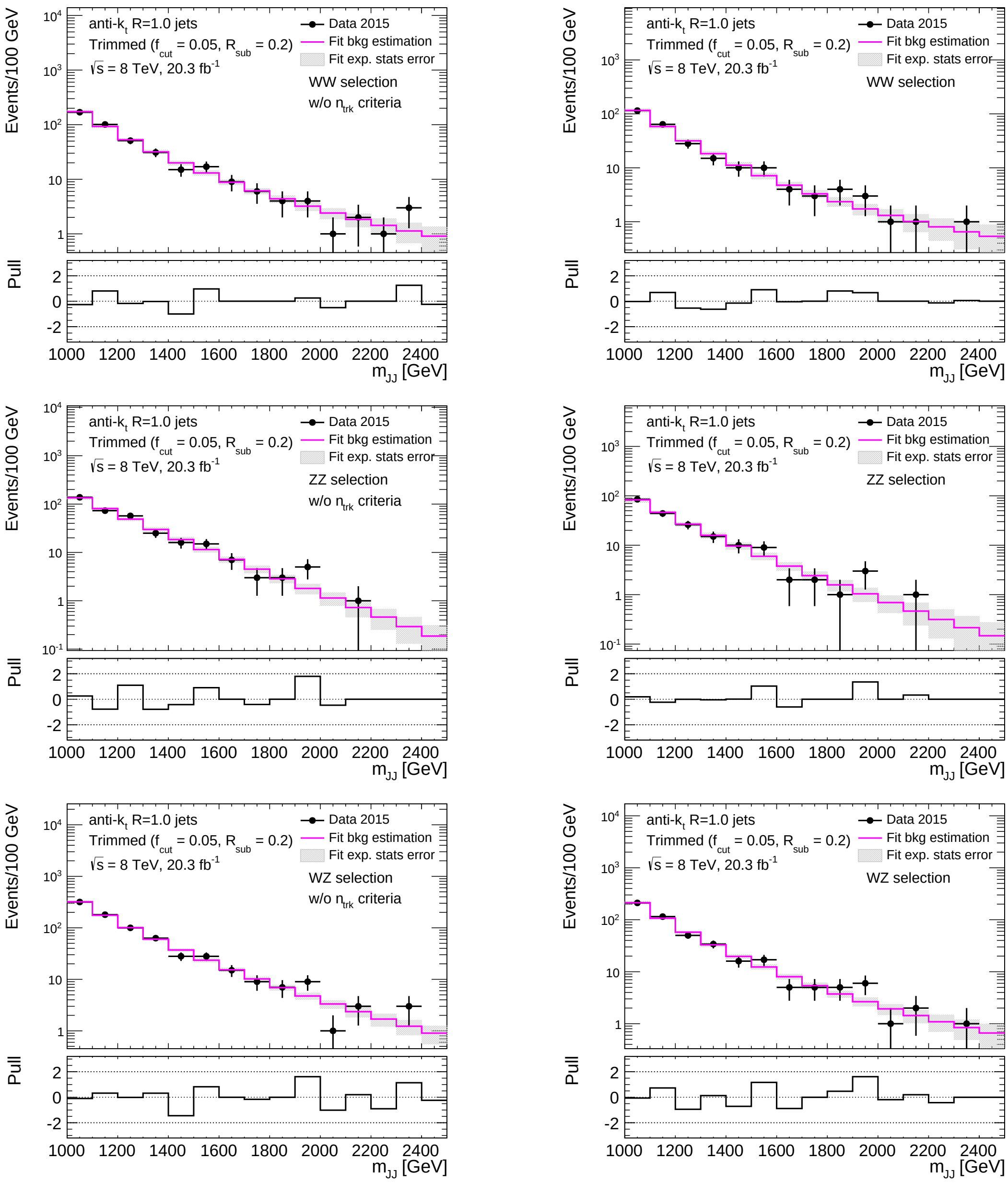


FIGURE 2.8: Background only-fit to the invariant dijet mass (m_{JJ}) distribution in $\sqrt{s}$ = 8 TeV data for boson-tagged jets using the Run-2 boson tagging algorithm described in Section 2.3 in the three channels *WW* (top), *ZZ* (middle) and *WZ* (bottom) without (left) and with (right) the requirement $n_{\text{trk}} < 30$ [170].

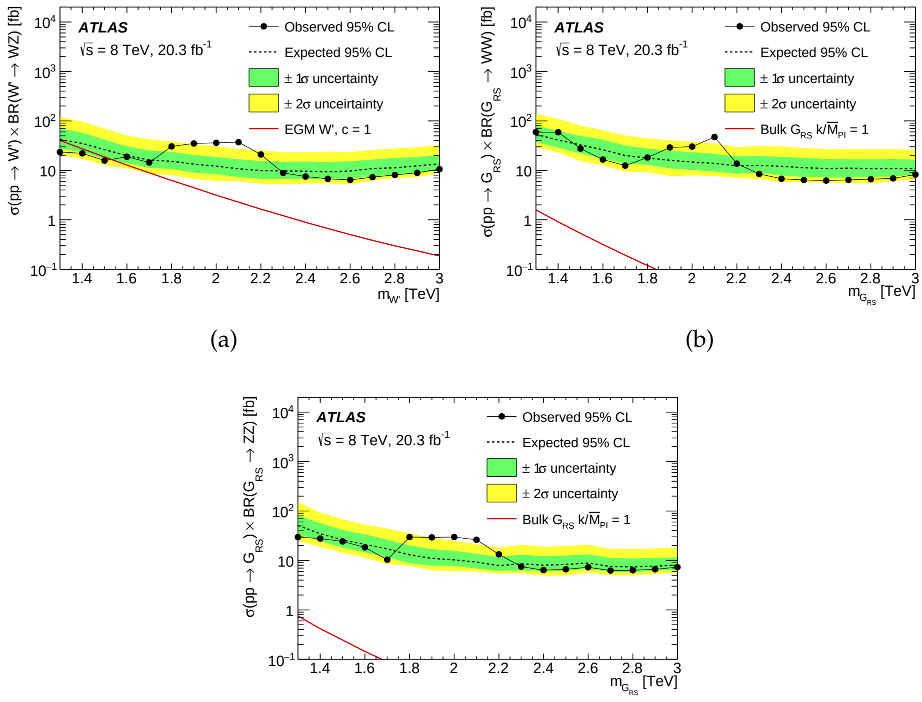


FIGURE 2.9: Upper limits at 95% CL on the cross-section times branching ratio of the (a) EGM W' model in the WZ signal region and the RS graviton model in the (b) WW and (c) ZZ signal regions. The green and yellow bands represent the 1σ and 2σ interval of the expected limit, respectively and the red line displays the predicted cross-section of the considered models [99].

Exclusion limits on new diboson resonances

In the absence of a significant deviation from the background-only hypothesis, quantified in terms of the local p_0 value, upper limits on the production cross-section times branching ratio of the different benchmark models can be set using the CLs method [211] as a function of the resonance mass using the EGM W' as a benchmark for the WZ selection, and the bulk G_{RS} model for the WW and ZZ selections. More detailed information on the statistical analysis can be found in Ref. [99]. Figure 2.9 shows the observed 95% CL upper limits on the cross section times branching ratio. A W' with EGM couplings and mass between 1.3 and 1.5 TeV is excluded at 95% CL. The cross section times branching ratio for excited graviton production with the model parameters mentioned in Section 2.1.2 is too low to be excluded with the sensitivity of this measurement. The biggest systematic uncertainties in the analysis correspond to the efficiency of the track-multiplicity cut, the jet mass scale and resolution as well as the parton shower modelling.

2.2 Combination with Other Diboson Resonance Searches at 8 TeV

To improve the sensitivity to new diboson resonances, a combination of the different ATLAS diboson resonance searches performed in ATLAS using the full 2012 dataset, corresponding to an integrated luminosity of 20.3 fb^{-1}, was performed [21]. Four different analysis were combined: the $\ell\nu\ell'\ell'$ [110], $\ell\ell qq$ [106], $\ell\nu qq$ [103] ($\ell = e,\ \mu$) and the fully hadronic final state presented in this Chapter [100]. The signal regions were completely decoupled as the

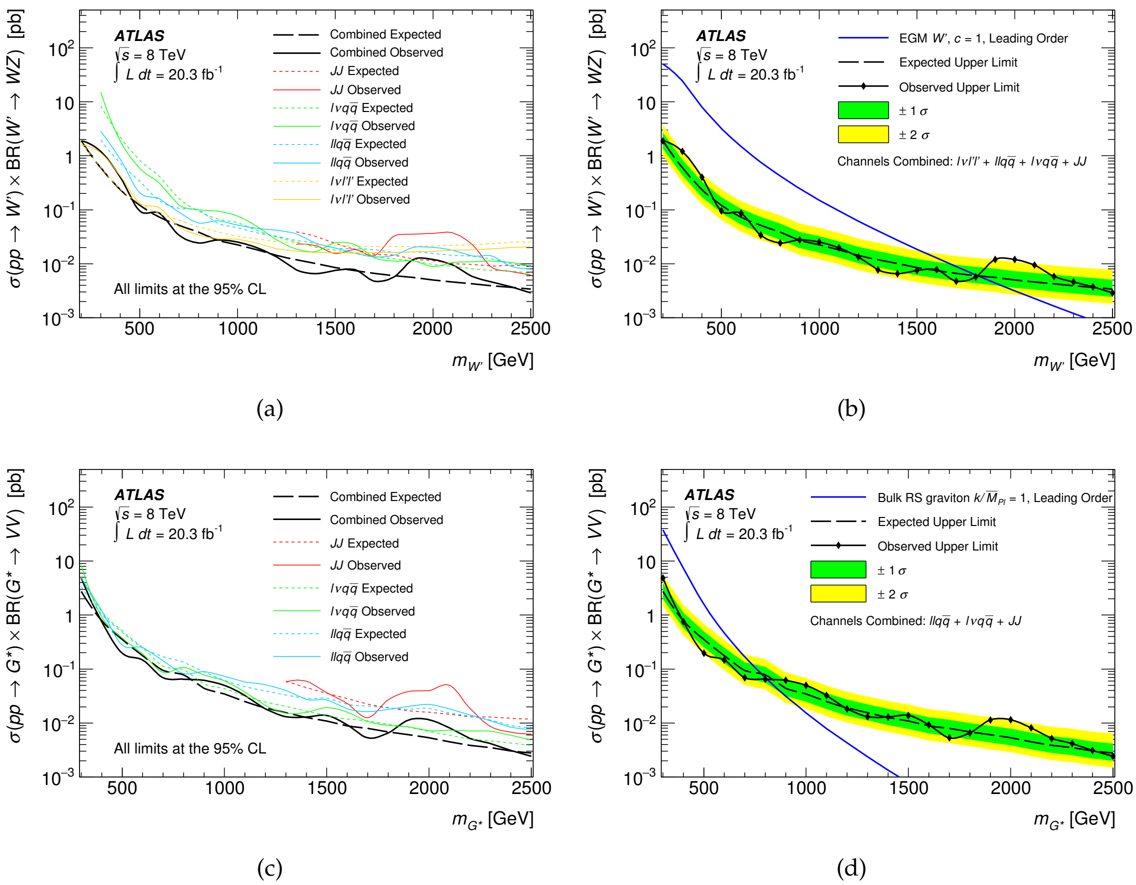
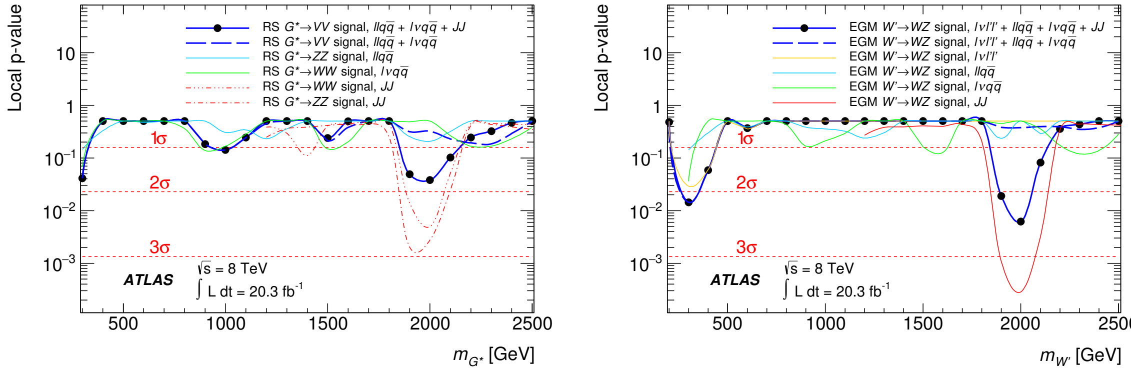


FIGURE 2.10: The expected and observed upper limits on the cross-section times branching ratio of the RS graviton model at 95% CL for the individual $\ell\ell qq$, $\ell\nu qq$ and $qqqq$ channels and their combination (left), and the comparison of the theory prediction to the 95% CL upper cross-section limits of the combined search (right). The green and yellow bands on the right represent the 1σ and 2σ interval of the expected limit, respectively and the blue line displays the predicted cross-section of the graviton model [21].

event selections were fully orthogonal. The semileptonic diboson channels define different signal regions with optimised selection criteria, referred to as low-p_{T} resolved/low mass (hadronic vector boson reconstructed two jets with $R = 0.4$) and high-p_{T} boosted/high mass (using a large-R jet to reconstruct the hadronic vector boson). The fully leptonic final state is only sensitive to low resonance masses whereas the semi-leptonic searches also improve the sensitivity to diboson resonances with high masses ($m_{JJ} > 1$ TeV) thanks to the boosted signal region included.

Figure 2.10 shows the observed 95% CL upper limits on the cross section times branching ratio for the individual channels as well as the combined diboson resonance search for the two benchmark signals used in Run-1. A W' with EGM couplings and mass below 1.8 is excluded at 95% CL, with an expected limit of 1.8 TeV. A graviton with $k/\bar{M}_{\mathrm{Pl}} = 1.0$ and mass below 810 GeV is also excluded, where a lower limit of 790 GeV was expected. The combination thus exceeds the lower mass limit of the most stringent individual analysis in the $\ell\nu qq$ final state by 50 GeV for W' and 0.22 TeV for the graviton.

As mentioned in Section 2.1.1 CMS also performed similar searches using the Run-1 dataset with comparable sensitivity to the ATLAS analyses, some of them included also



(a) (b)

FIGURE 2.11: The local p_0 value, obtained in the search for the G_{RS} (a) and EGM W' (b), as a function of the resonance mass for the individual searches and the combined diboson resonance search. Furthermore, the local significance is indicated by the red dashed horizontal lines [21].

diboson topologies with a Higgs boson decaying to two b-quarks. CMS analyses used a different boosted boson tagging strategy: C/A jets with $R = 0.8$, pruned and a requirement on the N-subjettiness variable τ_{21}. Figure 2.12 compares the results obtained in different final states. The limits of the different searches are marginally correlated and they used a different signal benchmark, a heavy vector triplet (HVT) based on a simplified phenomenological Lagrangian [206].

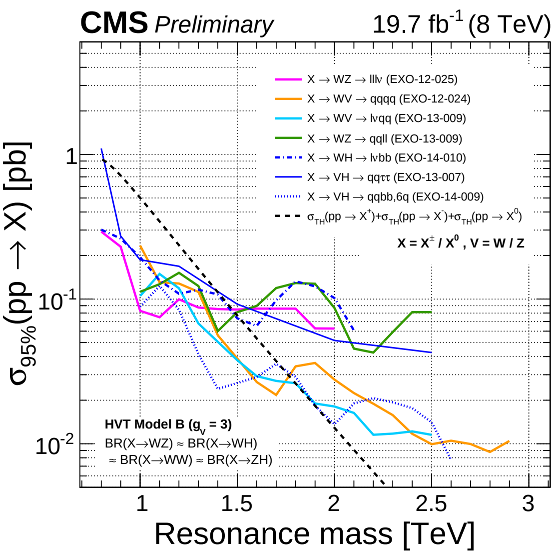


FIGURE 2.12: Observed 95% CL upper limits on the production cross section of a neutral (X^0) and charged ($X^{\pm}$) narrow spin-1 resonances decaying into a pair of bosons (VV or VH with $V = W/Z$). The curves represent the comparison of the results obtained in different final states and based on CMS data corresponding to an integrated luminosity of 19.7/fb recorded in proton-proton collisions at $\sqrt{s} = 8$ TeV. The black dashed curve represents the theoretical prediction for the sum of the production cross sections of the three spin-1 resonances degenerate in mass in the HVT model [161].

2.3 Preparation for Run-2: a more performant boosted W/Z boson tagger

2.3.1 Boosted boson identifications at $\sqrt{s} = 8$ TeV

At the time of the analysis described earlier in this Chapter a proper optimisation and comparison study was not available for boosted boson identification. Such study took place between 2014 and beginning of 2015 for $\sqrt{s} = 8$ TeV and was documented in Ref. [40], where more details can be found. This section briefly summarises that effort. A two-stage MC-based optimisation procedure was carried out to pair requirements on the signal groomed jet mass and substructure variables, providing a 50% efficiency for identifying W bosons with transverse momenta larger than 200 GeV while maintaining the multijet background efficiencies of 2-4% for jets with the same transverse momentum. Afterwards, these signal and background efficiencies were confirmed in data for a chosen selection of tagging techniques. Only the optimisation of W boson identification has been studied, however the techniques can be applied as well to identify Z bosons.

In the first step of the MC-based optimisation, 27 jet collections were selected out of more than 500 configurations on the pruning, trimming and split-filtering techniques. A configuration is defined by a jet reconstruction algorithm, the radius parameter R, the grooming technique and the chosen parameters of the grooming algorithm. The selection was based in two criteria: (i) low background efficiency within a mass window, containing 68% of the signal events, defined around the W boson mass and (ii) low pile-up impact based on the correlation between average jet mass and number of primary vertices. Studies were performed in three jet p_{T} bins: [200, 350), [350, 500) and [500, 1000) GeV. Jet configurations for which the 68% window does not contain the W boson mass, the signal mass distribution is strongly non-Gaussian or the background mass distribution has an irregular shape were not considered. These configurations would indicate that the W boson decay is not fully contained within the large-R jet and/or make it impossible to derive a jet mass calibration. In general, the performances of the best few configurations of trimming, pruning and split-filtering are similar at the defined 68% working point. It was also observed that the average masses of jets with larger radii have a more pronounced pileup dependence, and the trimmed jet mass has a weaker pileup dependence than that obtained with the pruning and split-filtering algorithms. For all jet algorithms, the pileup dependence is much reduced for groomed jets.

In the second stage, for the 27 jet collections selected, different substructure variables were investigated within the mass window containing 68% of the signal around the W boson mass and two different signal efficiency working points (50% and 25%) were defined. 26 different substructure variables have been studied in combination with the 27 grooming configurations to find the pairwise combinations that result in the smallest background efficiency for the two defined working points. Figure 2.13 shows the inverse of the background efficiencies for the 50% efficiency working point for the range $350 < p_{\mathrm{T}} < 500$ GeV.

Based on this initial MC-based optimisation 4 jet configurations with three different substructure variables, C_2, D_2 and τ_{21}^{wta}, were found to be particularly good in identifying boosted hadronically decaying W bosons in all p_{T} ranges considered in the studies:

- anti-k_t $R = 1.0$ trimmed jets with $f_{\mathrm{cut}} = 5\%$, $R_{\mathrm{sub}} = 0.2$,
- anti-k_t $R = 1.0$ trimmed jets with $f_{\mathrm{cut}} = 5\%$, $R_{\mathrm{sub}} = 0.3$,
- C/A $R = 1.0$ pruned jets with $Z_{\mathrm{cut}} = 0.15$, $R_{\mathrm{cut}} = 0.5$,
- C/A $R = 1.2$ split-filtered jet with $y_{\mathrm{cut}} = 0.15$, $R_{\mathrm{sub}} = 0.3$.

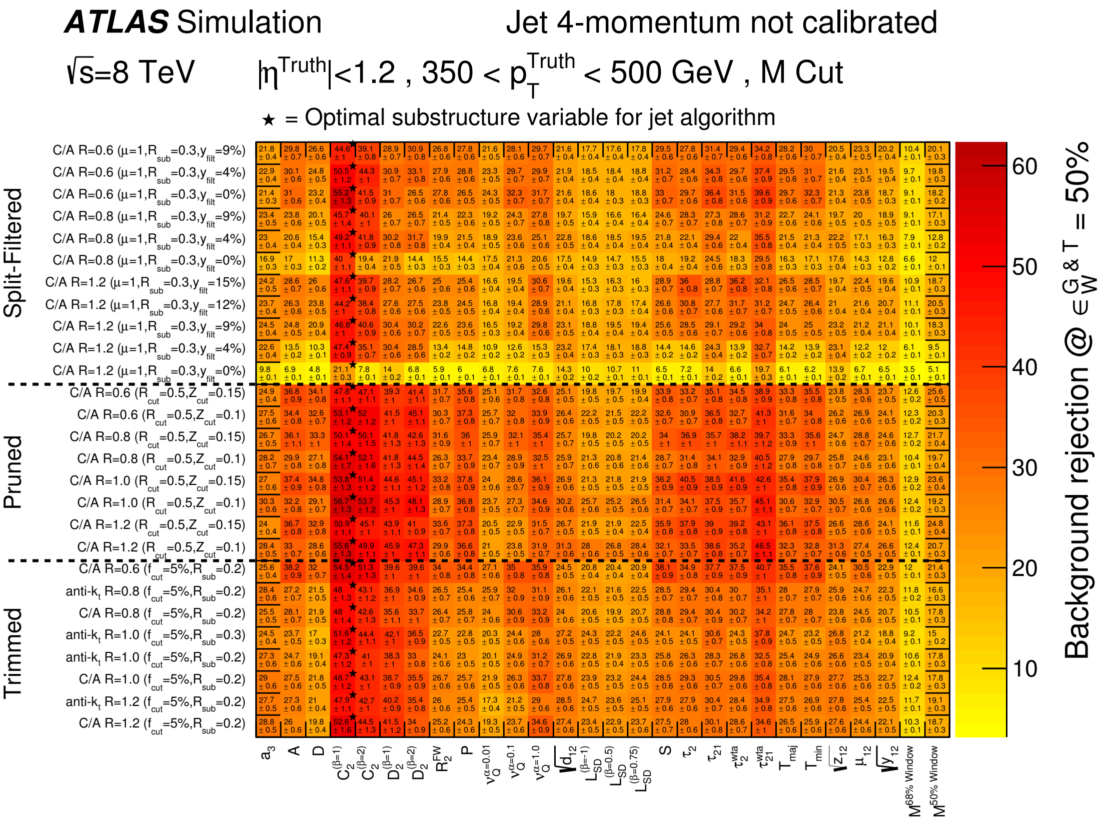
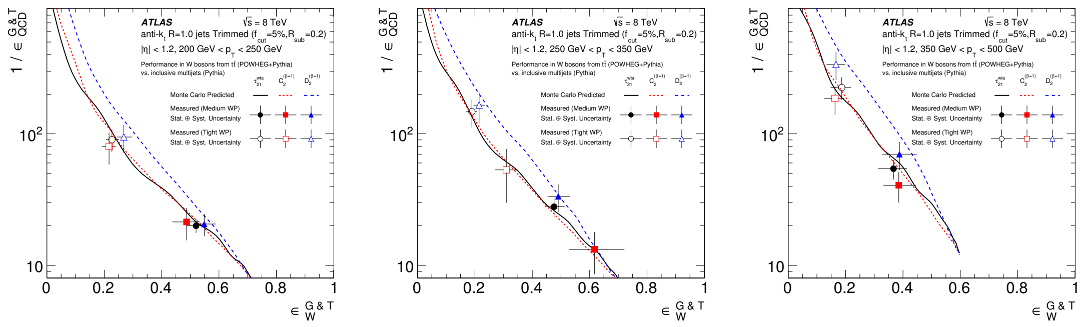


FIGURE 2.13: Background rejection corresponding to a signal efficiency of 50% for all possible combinations between the 27 grooming configurations and 26 substructure variables, after applying the uncalibrated groomed mass window requirement that provides a 68% signal efficiency. Jets with $350 < p_{\mathrm{T}}^{\mathrm{Truth}} < 500$ GeV are considered. The error shown are the result of the finite Monte Carlo sample size [40].



(a) (b) (c)

FIGURE 2.14: W boson tagging efficiency versus background rejection for jets with (a) $200 < p_{\mathrm{T}} < 250$ GeV (b) $250 < p_{\mathrm{T}} < 350$ GeV and (c) $350 < p_{\mathrm{T}} < 500$ GeV for C_2, D_2 and τ_{21}^{wta} after applying the 68% signal efficiency mass window [40]. The curves are obtained using MC simulation and compared to the 50% and 25% signal efficiency working point measurements in data using a lepton-plus-jet enriched $t\bar{t}$ sample and multi-jet events.

Since the first jet collection in the list was also found to be the only one of the four with negligible pileup dependence across all p_{T} ranges for the Run-1 data taking conditions, it was used in combination with the three substructure variables (C_2, D_2 and τ_{21}^{wta}) as a benchmark to measure the performance in data and to compare it to the predictions from MC simulations using the 20.3/fb of 8 TeV pp collisions recorded by ATLAS at the LHC in 2012. The W-tagging efficiency was measured using hadronically decaying W bosons are reconstructed from an enriched sample of top anti-top quark pair events in the lepton-plus-jets channel in data and in MC provided by Powheg-BOX + Pythia and MC@NLO + Herwig. The background efficiency is measured in multijet-enriched data and in Pythia and Herwig++ simulations. Details about the MC samples and selection used can be found in Ref. [40]. At this stage the jet mass and energy calibrations have been derived and applied, as described in Section 1.4.4.

The results of the signal efficiency versus background rejection curves in MC are shown in Figure 2.14. The data measurements are also shown for two efficiency working points, i.e. $\varepsilon_W = 50\%$ and 25%, defined by a combination of the 68% mass window requirements and a single-sided substructure variable cut. The signal efficiency was only measured in the range $200 < p_{\mathrm{T}} < 500$ GeV due to the lack of produced top quark pairs in data at high-p_{T}. The signal efficiency measured is not exactly 50% and 25% because the working points are calculated using W-jets from $W' \to WZ \to qqll$ where Ws are exclusively longitudinal polarised, and are measured in $t\bar{t}$ events where Ws have a longitudinal as well as transverse polarisation. Globally, the D_2 variable outperfoms C_2 and τ_{21}^{wta} and gives the largest background rejection for a fixed signal efficiency. In general the MC predictions show a good agreement within the uncertainties with the data measurements of signal efficiencies. In some configurations, significant differences are observed in both the signal and background efficiencies highlighting the potential for data measurements such as these to be used for tuning MC simulations.

This jet collection, anti-k_t $R = 1.0$ trimmed jets with $f_{\mathrm{cut}} = 5\%$ and $R_{\mathrm{sub}} = 0.2$, together with a selection on the jet mass and the D_2 variable was chosen as benchmark for further measurements in 8 TeV collision data.

2.3.2 Boosted boson identifications at $\sqrt{s} = 13$ TeV

The preliminary conclusions of the Run 1 studies summarised in Section 2.3.1 provided a subset of jet grooming algorithm configurations and substructure variables found to be efficient at identifying W-jets (from both exotic high-mass resonances and Standard Model $t\bar{t}$ production) whilst rejecting multijet and combinatorial backgrounds. The advent of Run-2 in 2015 required a similar optimization study taking into account the higher boost and higher pile-up conditions expected in 13 TeV data. Such study took place in 2015 and was documented in Ref. [39]. I was the contact physicist leading this effort and one of the editors of the public ATLAS note.

Jets groomed using each of the four configurations deemed promising based on previous studies, and three substructure variables (C_2, D_2, τ_{21}^{wta}) were studied:

R2: anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 5\%$, $R_{\text{sub}} = 0.2$. Benchmark jet collection from Run-1 optimisation studies [40].

P515: C/A $R = 1.0$ pruned jets with $Z_{\text{cut}} = 0.15$, $R_{\text{cut}} = 0.5$. This pruning configuration was found to be particularly good for tagging W-jets in the p_{T} 500 GeV region. It is similar to the algorithm favoured by CMS in Run 1 [150], using the same distance requirement of R_{cut} and a slightly harder momentum requirement $Z_{\text{cut}} = 0.15$ as opposed to the $Z_{\text{cut}} = 0.10$ favoured by CMS who also prefer an initial radius of $R = 0.8$.

Y15: C/A $R = 1.2$ split-filtered jets with $y_{\text{cut}} = 0.15$, $R_{\text{sub}} = 0.3$. The best-performing flavour of split-filtering from the extensive Run 1 studies [40] uses a relatively hard momentum balance cut of $y_{\text{cut}} = 0.15$ (compared to the 0.09 benchmark from Ref. [134]) and a fixed subjet radius $R_{\text{sub}} = 0.3$.

Y4: C/A $R = 1.2$ split-filtered jets with $y_{\text{cut}} = 0.04$, $R_{\text{sub}} = 0.3$. This split-filtering configuration was optimised specifically for the Run 1 search for hadronically decaying exotic diboson resonances described earlier in this Chapter [99].

Figure 2.15 shows the jet mass pile-up dependence for the four grooming configurations studied. It is possible to remove approximately all pileup effects using the R2 algorithm, with Y15 also performing very well. These observations applied to all the p_{T} ranges studied here.

The Run-2 optimisation study closely followed the strategy in the Run-1 optimisation study: the smallest mass window that contains 68% of the signal events is calculated and then a 50% and a 25% signal efficiency working points are defined by combining the 68% mass window and a single-sided cut on the C_2, D_2 and τ_{21}^{wta} substructure variables. Contrary to the Run-1 optimisation study, both the identification of W and Z bosons were optimised and a larger p_{T} spectra was studied: $200 < p_{\text{T}} < 2000$ GeV. The background rejection factors for the 50% signal efficiency working point are shown in Figure 2.16 separately for W and Z boson tagging for a low and a high-p_{T} bin. In the lower p_{T} range the mass-only requirement yields higher background rejection power for Z-jets than for W-jets. This is a result of the Z boson mass being higher and therefore better separated from the multijet background. Conversely, all of the pairwise combinations result in higher background rejection power for W-jets than for Z-jets. This behaviour is coming from the observed broader Z-jet mass distribution rather than from differences in the distributions of the substructure variables. There are a few combinations that perform particularly well in terms of background rejection, namely P515 and R2 with C_2 and D_2. Taking into account that the pruned jet mass exhibit a non-negligible pile-up dependence at low-p_{T}, the R2 D_2 combination was chosen as the best option for a baseline Run 2 boson tagger.

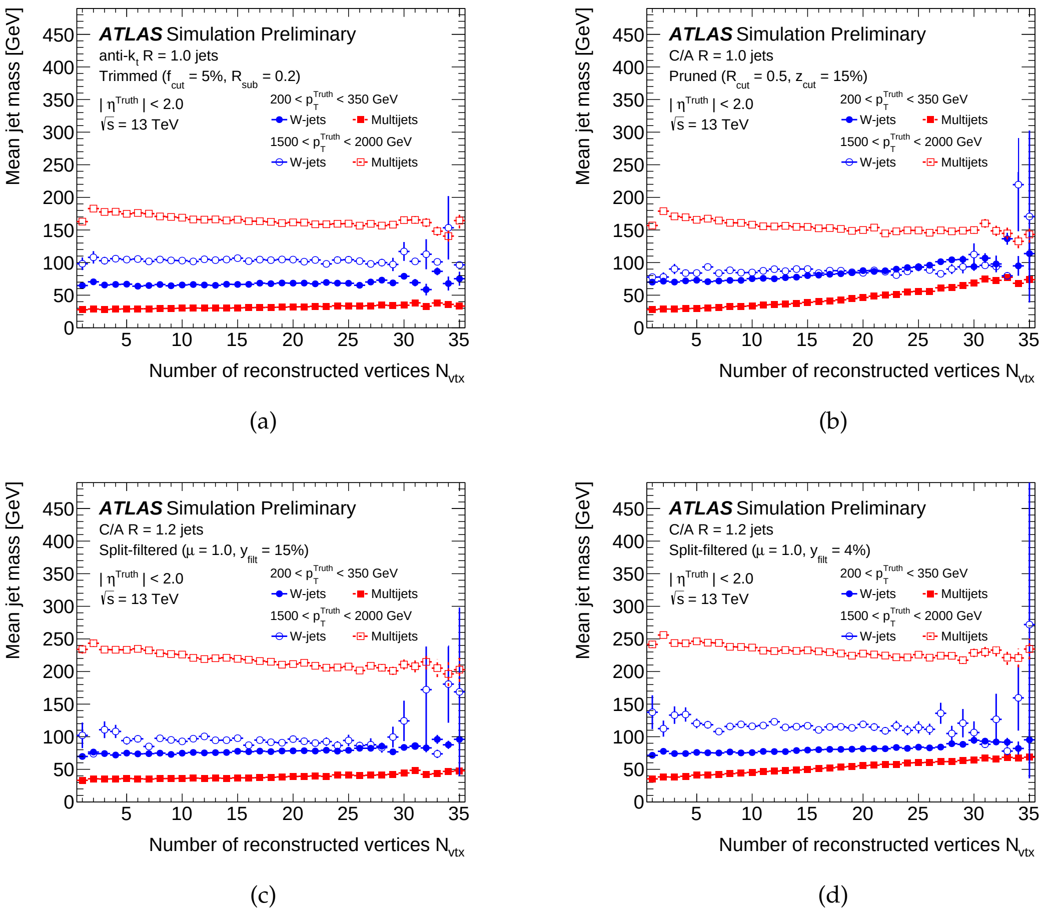


FIGURE 2.15: The pile-up dependence of the mean jet mass $\langle M \rangle$ for groomed signal W-jets (blue circles) and background jet (red squares) for two p_{T} ranges. The performance of the groomers in terms of removing the NPV-dependence for Z-jets is very similar to that for W-jets, so is not shown here for clarity [39].

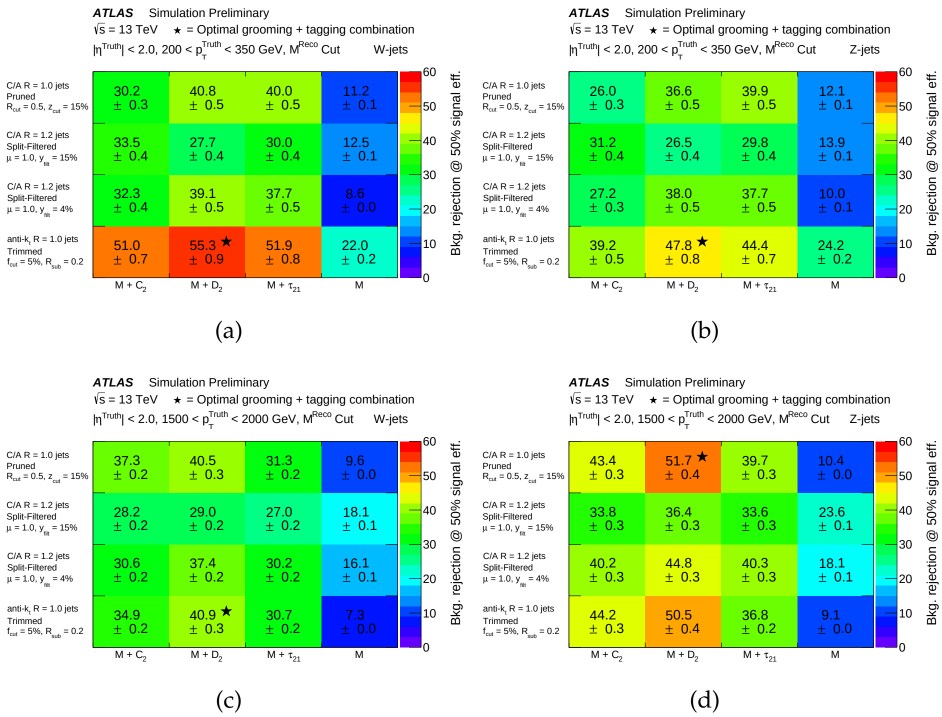
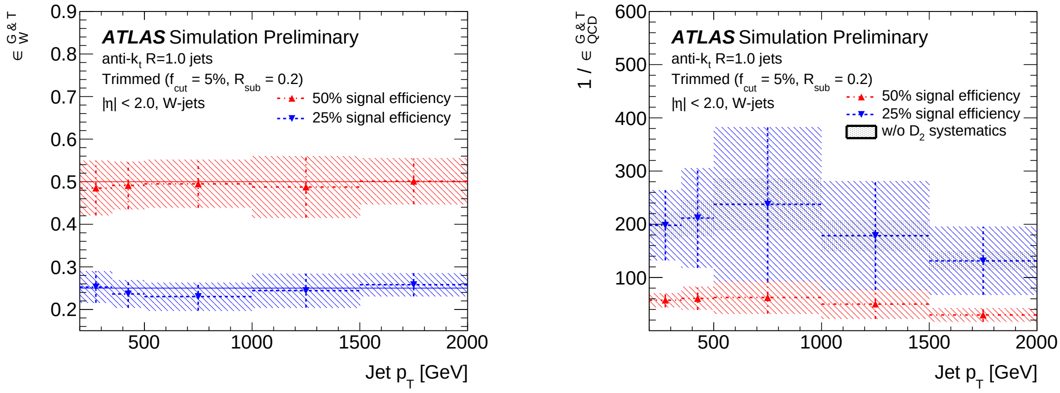


FIGURE 2.16: Matrix showing the background rejection achievable when tagging W-jets and Z-jets with $\epsilon^{\mathrm{G\&T}}_{W/Z} = 50\%$ with each pairwise combination of variables and for mass-only, for the four algorithm configurations in the range $200 < p_{\mathrm{T}} < 350\mathrm{GeV}$ and $1500 < p_{\mathrm{T}} < 2000\mathrm{GeV}$ [39].



(a) (b)

FIGURE 2.17: (a) Signal efficiency for W-jets and (b) background rejection in QCD dijet samples for two different working points with $\varepsilon_{\text{sig}}^{\text{G\&T}} = 50\%$ and 25% [39].

The baseline vector boson identification algorithm was then further optimised to be used in physics analyses. Dedicated jet energy and mass calibrations are derived, using the procedure described in Section 1.4.4. Smooths cuts as a function of p_{T} on the jet mass and D_2 were derived to obtain the 50% and a 25% signal efficiency working points.

Figure 2.17 shows the signal efficiencies and background rejection for W-tagger optimised and recommended for physics analysis at the beginning of Run-2. Similar results were found for the Z-tagger. The uncertainty bands include uncertainties on the jet p_{T}, mass and D_2 scale, derived for early Run-2 analysis described in Section 1.4.4 as well the corresponding resolution uncertainties. Detailed information about the samples used to derive those uncertainties can be found in Ref. [39]. The large uncertainties on the background rejection for the tight working point are dominated by the D_2 scale uncertainties. For a signal efficiency of 50%± 5%, a background rejection of 60± 10 can be achieved.

2.4 Summary

From all the great results released by the LHC during the Run-1 data taking period, the search for a heavy resonance decaying hadronically to dibosons stood up. The ATLAS Run-1 data showed an excess at 2 TeV, with a local significance of 3.4 sigma and a 2.5 sigma global significance. Other searches in ATLAS looking for diboson resonances in other channels (including leptonic and semileptonic diboson decays) showed no excess. However, at the other side of the LHC ring, CMS Run-1 data seemed to show a small excess as well at the same place. Expectations for 2015 were high. The increased LHC beam energy in 2015 at the start of Run-2 meant that, if existing, heavy resonances would be made much more easily and next chapter will describe what happened with this search at the beginning of Run-2.

At previous hadronic colliders, the decay of W and Z bosons to quarks was quite difficult to observe over the background of QCD jets produced directly in the proton collisions. But at the LHC, hadronic boson reconstruction has become a standard tool. In this chapter we described the first large boosted boson tagger optimisation efforts that took place in ATLAS at the end of Run-1 and updates performed at the beginning of Run-2. During Run-2 the simple two-variable boosted W/Z tagger described earlier, made up of a groomed mass window and a D_2 requirement, was widely used by ATLAS physics analyses, including the diboson resonance searches that will be presented in the next Chapter 3.

New optimisations were made in the last couple of years for new jet collections like TCC [91, 164] and UFO [87, 164] leading to an improved background rejection thanks to the good angular resolution of UFOs and TCCs. Recent extensions to these studies include combining more variables (like the number of track associated to the large-R jet), using multivariate neural network techniques and investigating the use of new and more performant pileup-removal techniques like constituent subtraction + SoftKiller (CS+SK) combined with soft-drop grooming [87]. All these developments has greatly improved the sensitivity of physics analyses using the full Run-2 dataset and are a solid baseline for Run-3.

Chapter 3

Searching for diboson resonances in Run-2

After the observation of an excess of events over the background prediction with a global significance of 2.5σ in the VV diboson resonance search at 8 TeV(see Section 2.1), repeating the analysis using Run-2 $\sqrt{s} = 13$ TeV data was of uppermost interest in the scientific community. Due to the increase of parton luminosity from 8 to 13 TeV, new resonances with high masses, initiated by gg, qg or $q\bar{q}$ production, were produced more abundantly and therefore Run-2 data helped us understanding the nature of the excess observed.

No excess was seen with the early Run-2 data collected. The analysis was redone with the first 3.2/fb of Run-2 data, accumulated in 2015 [104, 111]; and I was the analysis team coordinator. I also worked on the analysis of the 36.1/fb of data available at the end of 2016, also the subject of a publication in 2017 [92]. In parallel to this analysis work, I was in charge of the Diboson/Multilepton subgroup of the ATLAS exotic group for one year between April 2016 and April 2017, which involved the coordination of about 15 analyses and close work with theorists in order to define the theoretical benchmarks to be used in these searches. The analysis strategy broadly followed the one presented in the previous Chapter. The main differences with respect to the 8 TeV analysis and results obtained are detailed in Section 3.2. My main contributions were the coordination of the working group of 30 people, the background estimation and the boson jet identification studies already described in Section 2.3.2.

During this period I got also interested in expanding the hadronically decaying diboson searches to topologies including Higgs($\rightarrow b\bar{b}$) bosons. I contributed to the first VH hadronic resonance searches done in ATLAS [10, 96], in particular with background estimation studies and the boosted Higgs boson identification studies briefly described in Section 1.4.5. The hadronic VH and VV searches results using 36.1/fb of data available at the end of 2016 were combined with their leptonic counterpart and the dilepton resonance searches. This work, for which I was one of the analysis coordinators, was also subject of a publication [19] and the main results are found in Section 3.3.

3.1 Theoretical models used

The results presented in this Chapter are interpreted in the context of three models: the Randall-Sundrum (RS) model, an empirical model featuring a new heavy scalar and the heavy vector triplet (HVT) model based on a simplified phenomenological Lagrangian [206] which replaces the extended gauge model W' used in Run-1 searches.

The RS model used is the same described in Chapter 2 with $k/\bar{M}_{\rm Pl} = 1$, produced mainly via gluon-gluon fusion (ggF) and for which the Kaluza–Klein excitations of gravitons, G_{RS} (corresponding to the resonance we were looking for), width relative to its mass is approximately 6% [209, 210]. The second model considered was an empirical model with a narrow heavy scalar resonance produced via ggF and vector boson fusion (VBF) mechanisms and decaying directly into VV. The width of this new scalar was assumed to be negligible compared with the detector resolution, and the relative branching fractions for decay into the WW and ZZ final states approximately follow a 2:1 ratio. This benchmark was used to explore the analysis sensitivity to extended Higgs sectors.

The HVT model provides a broad phenomenological framework including different scenarios involving new heavy gauge bosons and their couplings to SM fermions and bosons. A new vector triplet is introduced with one charged (W') and one neutral (Z') heavy spin-one particle which mixes with the SM vector bosons and fermions. The W' and Z' states are collectively denoted by V'. The coupling of the new vector triplet to the fermions is given by $g_f = (g^2/g_V^2) \cdot c_{\mathrm{F}}$, where g is the coupling constant of the $\mathrm{SU}(2)_L$ gauge group of the weak interaction, g_V the coupling strength to the new vector boson and c_{F} is a free parameter of the order of one. Therefore, couplings specific to quarks and leptons are given by $g_q = g^2 c_q / g_V$ and $g_\ell = g^2 c_\ell / g_V$. The coupling to the Higgs boson is given by $g_H = g_V c_H$.

Three HVT scenarios are used. The first two benchmarks are both Drell-Yan (DY) production mechanisms, while the third benchmark proceeds via the vector-boson fusion (VBF) mechanism:

- Model A, the first DY scenario, reproduces the phenomenology of weakly coupled models based on an extended gauge symmetry [125]. In this case, the couplings are $g_H = -0.56$ and $g_f = -0.55$, with the universal fermion coupling $g_f = g_q = g_\ell$ and $g_V = 1$. The width of the W' and Z' in this model is approximately 2.5% of the resonance mass.

- Model B, the second DY scenario, implements a strongly coupled scenario as in composite Higgs models [169] with $g_H = -2.9$ and $g_f = 0.14$ and $g_V = 3$. The V' resonances are broader than in the weakly coupled scenario, model A, but remain narrow (below 5%) relative to the experimental resolution.

- Model C, focuses only on vector-boson fusion processes. In this case, the V' resonance couplings are set to $g_H = 1$ and $g_f = 0$. Model C is therefore in a separate phase space domain to models A and B and assumes no DY production.

For the DY process with decay of the V' into lepton–antilepton final states, branching fractions are largest in model A with values of approximately 4% and only about 0.2% in model B, for each generation taken separately. In contrast, the branching fractions for decays into individual diboson channels are about 2% in model A, whereas they are close to 50% in model B.

The interpretation of diboson and dilepton searches can be extended beyond these three benchmark models by considering the two-dimensional parameter space consisting of g_H and g_f (assuming fermion universality) or g_q and g_ℓ for a given value of g_H. The different production mechanisms and decay modes included in Section 3.3 provide sensitivity to different regions of this parameter space, with production via the DY process providing sensitivity to g_q and production via VBF providing sensitivity to g_H. Likewise, decays into lepton–antilepton states provide sensitivity to g_ℓ whereas decays into diboson states provide sensitivity to g_H [19].

Table 3.1 presents production cross sections for several heavy resonance masses in the HVT models A, B, and C, and the bulk RS model.

TABLE 3.1: Cross sections for production of heavy resonances of different masses in HVT models A and B via the Drell–Yan process, in HVT model C via vector-boson fusion, and in the bulk RS model via gluon–gluon fusion and the Drell–Yan process [19].

	HVT model A		HVT model B		HVT model C		Bulk RS
m [TeV]	$\sigma(W')$ [fb]	$\sigma(Z')$ [fb]	$\sigma(W')$ [fb]	$\sigma(Z')$ [fb]	$\sigma(W')$ [fb]	$\sigma(Z')$ [fb]	$\sigma(G_{RS})$ [fb]
1.0	2.20×10^4	1.12×10^4	987	510	1.30	0.888	583
2.6	219	100	14.0	6.44	4.78×10^{-3}	3.14×10^{-3}	1.41
4.0	9.49	4.37	0.626	0.288	1.27×10^{-4}	7.92×10^{-5}	3.25×10^{-2}

3.2 Hadronically VV resonance search at 13 TeV

3.2.1 Using 3.2/fb accumulated in 2015

Changes with respect to Run-1 analysis

The very first data at 13 TeV were analyzed. Although the initial 3.2/fb data collected at the beginning of Run-2 are only a fraction of the integrated luminosity analyzed in the Run-1 publication, the increase of the energy at the center of mass allowed to increase the sensitivity of the analysis. This section is largely based on Ref. [104, 111]. The main changes with respect to the Run-1 hadronically VV resonance search described in Section 2.1 are:

- The Heavy Vector Triplet model A with $W' \to WZ$ and $Z' \to WW$ is used for the results interpretation, instead of the extended gauge model.
- The boosted boson tagger was optimized to account for the higher boost and changes in the pile-up conditions as described in Section 2.3.2. Table 3.2 summarizes the differences between the boosted boson tagger used in Run-1 and Run-2.
- The trigger used was the same, single-jet trigger with a nominal p_T threshold of 360 GeV. However, due to a smaller jet radius side used in Run-2 the leading large-R p_T (m_{JJ}) threshold could be decreased from 540 (1050) to 450 (1000) GeV.
- The analysis was still dominated by the uncertainties on the scale and resolution of the large-R jet observables: the jet p_T, mass and D_2. But there were changes in the estimation of the uncertainties used:
  - The scale uncertainties are larger than the ones of the Run-1 analysis since they are based on the Run-1 uncertainties with additional MC-based uncertainties to account for differences in the detector and jet reconstruction between 8 TeV and 13 TeV as described in Section 1.4.4. The uncertainty on the jet p_T scale changed from 2% to 5%, on the jet mass scale from 3% to 6% and for the D_2 variable is 4-5% instead of 2% for $\sqrt{y_{12}}$.
  - The uncertainty associated to the n_{trk} jet cut was reduced from 20% to 6% because of the better modelling of the n_{trk} distribution in Run-2.
  - The uncertainty on the luminosity estimation increased from 2.8% to 5%.

	Run-1	Run-2
Jet algorithm	C/A $R = 1.2$	anti-k_t $R = 1.0$
Jet grooming	Split-Filtering:	Trimming:
	$R_{\text{sub}} = 0.3$, $y_{\text{cut}} = 0.04$, $\mu = 100\%$	$f_{\text{cut}} = 5\%$, $R_{\text{sub}} = 0.2$
Jet tagging	Mass window ± 13 GeV	Mass window ± 15 GeV
	$\sqrt{y_{12}} \geq 0.45$	D_2 p_{T}-dependent cut
	$n_{\text{trk}} < 30$	$n_{\text{trk}} < 30$

TABLE 3.2: Summary of the Run-1 and early Run-2 boosted boson tagging algorithms used in the hadronically VV exotics resonance search analysis [170].

Results and exclusion limits obtained

Given the Run-1 results, both Run-1 and Run-2 boosted boson taggers were compared in this version of the analysis. For low resonance masses, higher signal efficiencies were obtained with the Run-1 tagger than the Run-2 tagger and vice-versa for high resonance masses. As mentioned in Section 2.3.2 the Run-2 tagger keeps an almost constant efficiency as a function of p_{T} by construction. Regarding the predicted background rejection, it was found to be similar for both taggers for high m_{JJ} while the Run-2 tagger outperforms the Run-1 tagger for low m_{JJ}. Figure 3.1 shows the fitted dijet mass distribution of events passing the WW, ZZ, WZ when using the Run-1 and the Run-2 boosted boson tagger. No significant deviations were observed in either case.

The total number of events observed in the official analysis, using the Run-2 boosted boson tagger, was 95, 77 and 128 in the WW, ZZ and WZ signal region, respectively. Due to the large-R jet mass resolution, the chosen mass window around the W boson and Z boson mass overlap and therefore the three signal regions are not orthogonal: 38 events are shared among all three signal regions. The fraction of events selected by both the WZ and WW selection are 60% whereas 56% of the events are common to the WZ and ZZ selection. When using the Run-1 boosted boson tagger a total of 259, 199 and 470 events are selected in the WW, ZZ and WZ signal regions, respectively. This is consistent with the observation of the expected background rejection being lower for the Run-1 tagger.

The data are in good agreement with the background expectations at all masses studied. Figure 3.2 shows the observed 95% CL upper limits on the cross section times branching ratio as a function of the resonance mass using the HVT and the bulk G_{RS} models. These results exclude the HVT W' model for masses between 1.38 and 1.6 TeV at the 95% CL. No exclusions of the other models considered were possible.

This early Run-2 version of the full hadronic diboson resonance search analysis was not sensitive enough to exclude the excess of events observed at m_{JJ} around 2 TeV in the Run-1 analysis. More statistics was needed to conclude on the observed excess as will be shown shortly in the next section. However, **the combination of the fully hadronic diboson resonance search with analyses in the semi-leptonic final state using 3.2 fb^{-1} of 13 TeV data in Ref. [111] exceeded the sensitivity of the 8 TeV analysis and at 2 TeV the limit obtained with the combination was very close to that expected from background. There was no sign of the 2 TeV peak around.**

3.2.2 Using 36.1/fb accumulated in 2015 and 2016

One year later, at the end of 2016, ten times the amount of data has already been accumulated compared to the first Run-2 release, providing ample justification for a re-release of

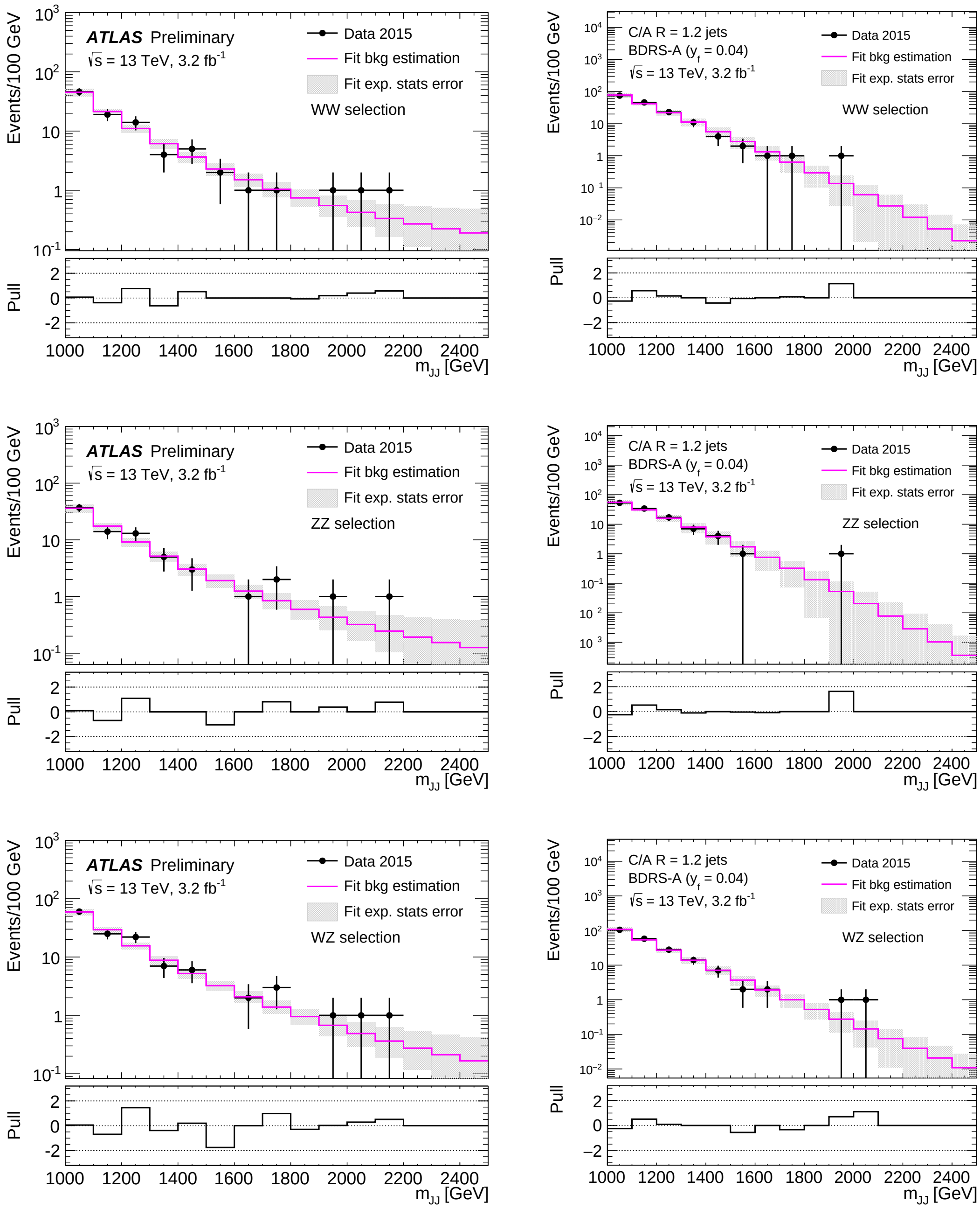


FIGURE 3.1: Background-only fit to the invariant dijet mass (m_{JJ}) distribution in 3.2/fb $\sqrt{s}$ = 13 TeV data for boson-tagged jets in the three different signal regions: (top) WW, (middle) ZZ, (bottom) WZ using the Run-1 (left) and Run-2 (right) boosted boson tagger [104, 170].

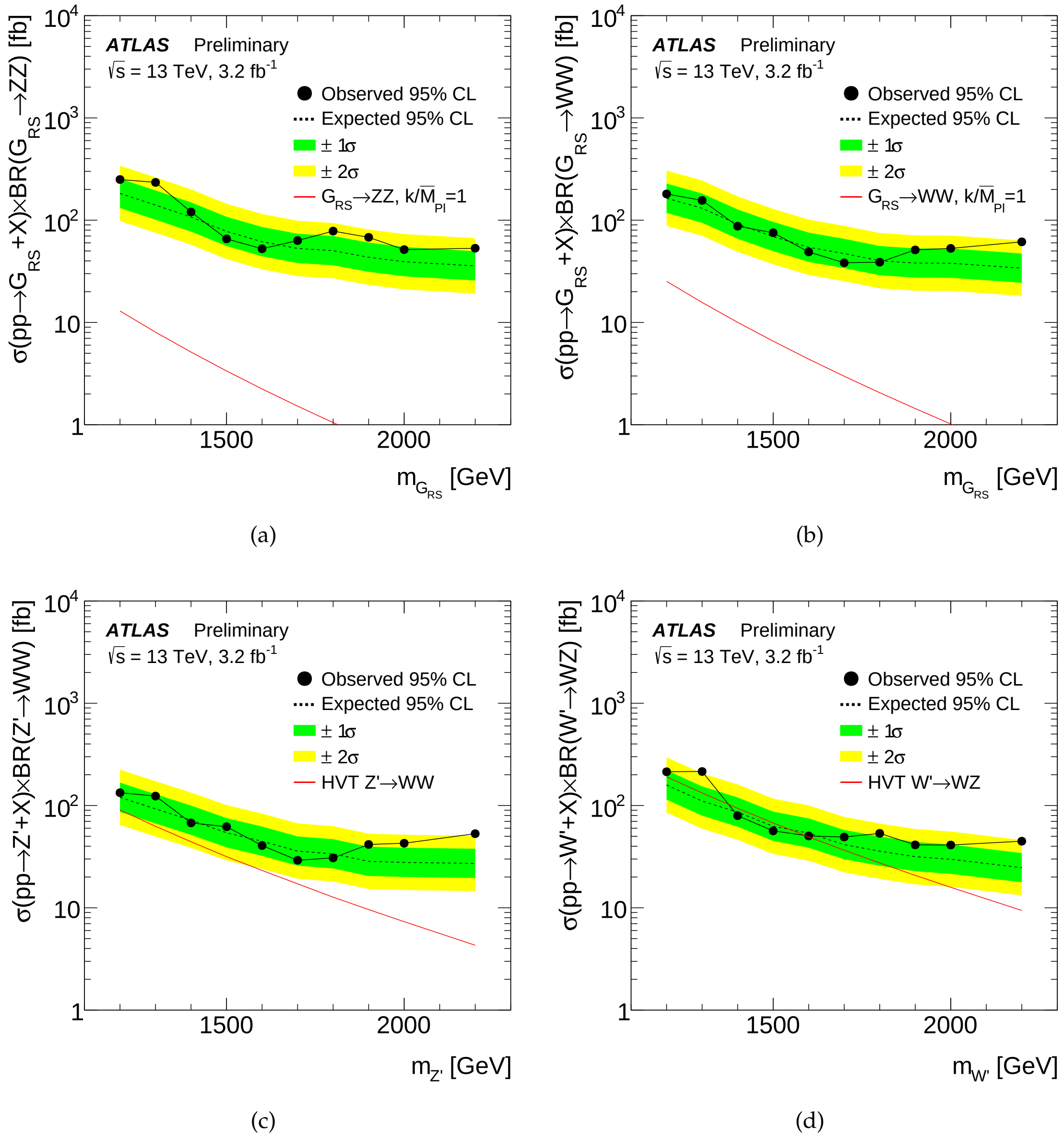


FIGURE 3.2: Upper limits at 95% CL on the cross-section times branching ratio of the bulk RS graviton through (a) the ZZ and (b) the WW channels, and on the HVT through (c) the $Z' \to WW$ and (d) the $W' \to WZ$ analyses. The green and yellow bands represent the 1σ and 2σ interval of the expected limit, respectively and the red line displays the predicted cross-section of the considered models [104].

the analysis. The main changes with respect to the 2015 hadronically VV resonance search described in the previous section were:

- A spin-0 signal was introduced for interpretation, as the heavy scalar resonance model (see Section 3.1). A heavy scalar is produced via gluon-gluon fusion with subsequent decay into a pair of vector bosons.
- Additional 33.5/fb of pp collisions from 2016 were analyzed. A trigger-level requirement of having at least one large-R jet with transverse energy above 360 GeV in 2015 and 420 GeV in 2016 was used.
- The invariant mass of the dijet system was raised to $m_{JJ} > 1.1$ TeV to avoid inefficiencies due to the minimum jet-p_{T} requirements and to guarantee that the trigger requirement is fully efficient.
- The combined jet mass, described in Section 1.4.5, was used instead of the calorimeter mass. The boosted boson tagger was re-optimized using the combined mass to achieve the maximal background-jet rejection for a fixed W or Z signal-jet efficiency of 50%. The background-jet rejection improved between 20-40% depending on the p_{T} range.
- The dominant sources of uncertainty are related large-R jet energy and mass calibrations. The main changes are:
  - JES, JMS and D_2 scale uncertainties were evaluated using the track-to-calorimeter double ratios between data and MC simulation (see Section 1.4.4 for more information). The size of the total correlated JES (JMS) uncertainty varies with jet p_{T} and is around 3% (5%) for the full signal mass range. The uncorrelated scale uncertainty in D_2 also varies with jet p_{T} and is approximately 3% per jet.
  - The luminosity estimation uncertainty decreased from 5% to 3.2%.

Figure 3.3 shows the resulting dijet mass distributions. A total of 497, 904, 618, 980, and 904 events are found in the WW, WZ, ZZ, $WW + WZ$, and $WW + ZZ$ signal regions. Approximately 20% of events are shared by all three regions: WW, WZ and ZZ. The requirements of the WW (ZZ) selection are satisfied by 47% (57%) of the events in the WZ signal region. No events are observed beyond 3.1 TeV. The dijet mass distributions in all signal regions are described well by the background model over the whole range explored.

Figure 3.4 shows the observed 95% CL upper limits on the cross section times branching ratio. These results exclude the HVT model A (model B) for masses between 1.2 and 3.1 TeV (1.2 and 3.5 TeV) at the 95% CL. For the bulk RS model a spin-2 Kaluza–Klein graviton is excluded in the range between 1.3 and 1.6 TeV. Upper limits on the production cross section times branching ratio for new heavy scalar particles are set with values of 9.7 fb and 3.5 fb at scalar masses of 2 TeV and 3 TeV, respectively. **This analysis allowed to exclude the excess of events observed at an invariant dijet mass of 2 TeV in the Run-1 analysis in the full hadronic VV topology alone.**

3.3 Combination of VV, VH and Leptonic Resonance Searches in Run-2

Between 2017 and 2018 I was one of the coordinators of an interesting effort aiming to combine for the first time in ATLAS the results of not only the VV and VH searches but lepton-antilepton searches as well. It used the most (at the time) recent ATLAS results obtained at 13 TeV with an integrated luminosity of approximately 36 fb^{-1}. This combination allowed

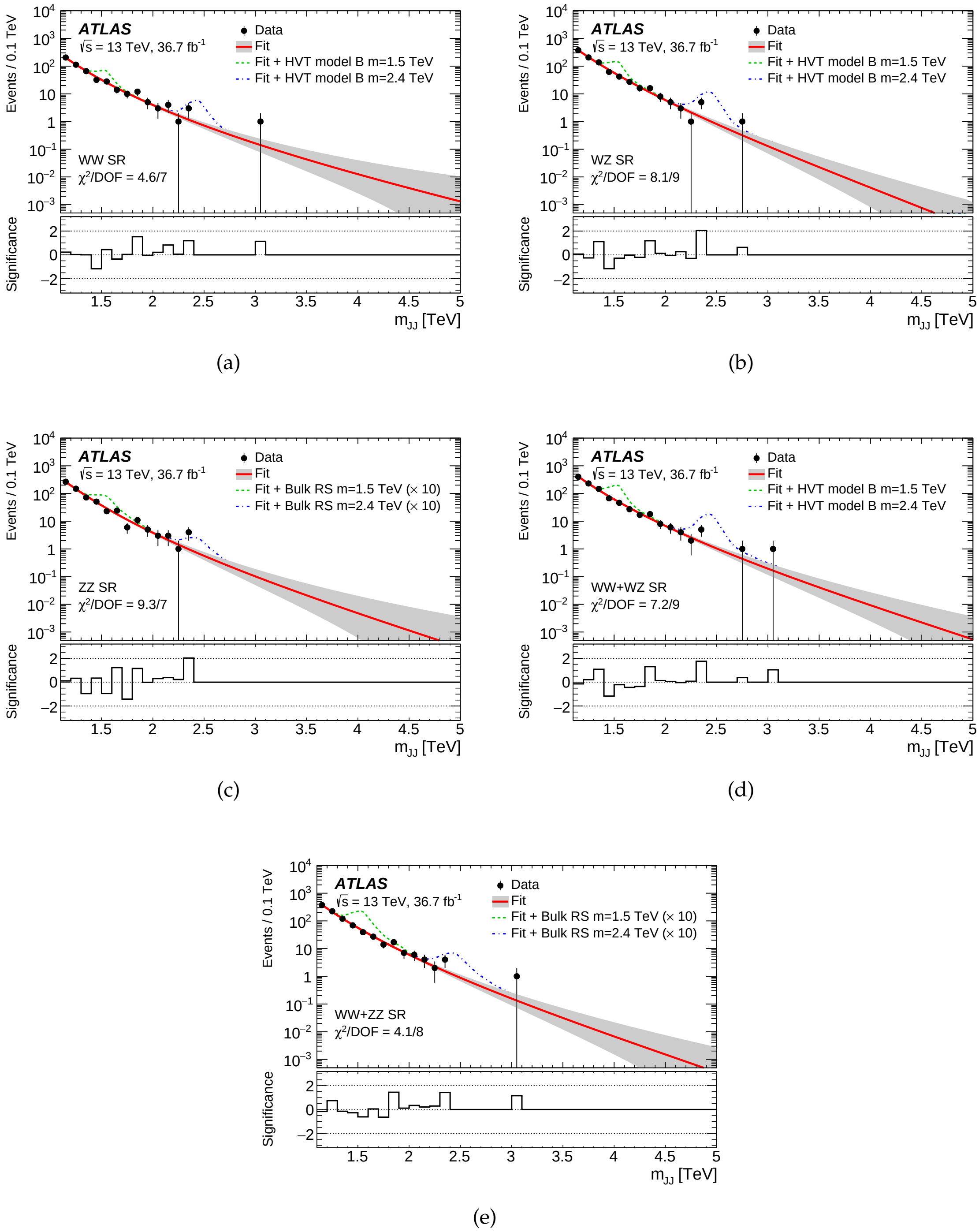
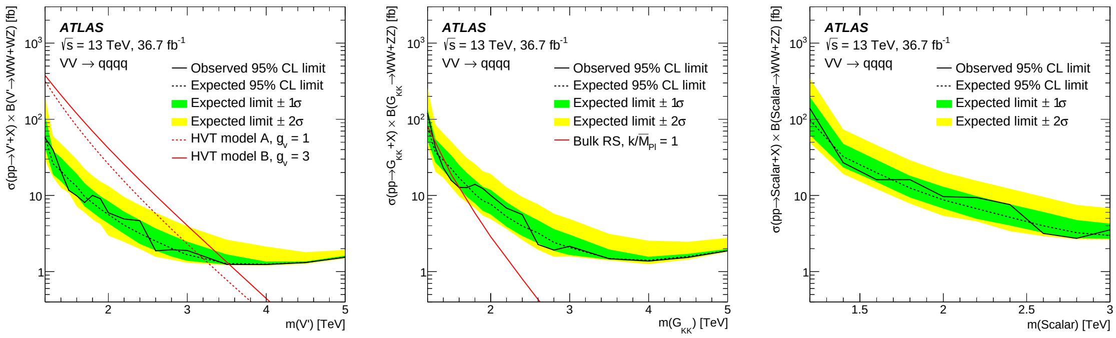


FIGURE 3.3: Background-only fit to the invariant dijet mass (m_{JJ}) distribution in 36.1/fb $\sqrt{s} = 13$ TeV data for boson-tagged jets in the three different signal regions: (a) WW, (b) WZ, and (c) ZZ signal regions, as well as in the combined (d) WW+WZ and (e) WW+ZZ signal regions. The lower panels show the significance of the observed event yield relative to the background fits [92].



(a) (b) (c)

FIGURE 3.4: Upper limits at 95% CL on the cross-section times branching ratio of the HVT through (a) the $V' \rightarrow WW + WZ$ production, (b) the bulk RS graviton through $WW + ZZ$ production and (c) the heavy scalar model through $WW + ZZ$ production. The green and yellow bands represent the 1σ and 2σ interval of the expected limit, respectively and the red line displays the predicted cross-section of the considered models [92].

to explore the complementarity of these searches and set stronger constraints over the 3 models of physics beyond the SM described in Section 3.1: HVT, Bulk RS and an empirical model featuring a new heavy scalar. This section is largely based on Ref. [19], where further details can be found.

The analyses generally searched for narrow resonances in the final-state mass distribution with the signal shape extracted from MC simulation of the specific models considered. The background shape and normalization were extracted from a combination of MC simulation and data, often relying on dedicated control regions to extract the various background contributions. Table 3.3 summarizes the analyses channels used in the combination, their experimental signature as well as the reference where the full description of each analysis can be found. Orthogonality between the different channels was a guiding criteria when defining the event selection.

TABLE 3.3: Summary of analysis channels, diboson states they are sensitive to, and their experimental signatures. The selection reflects requirements specific to each channel. Additional jets (not included in the "Jets" column) are required to define VBF categories. The notation "j" represents small-R jets for resolved/low boosted bosons and "J" represents large-R jets used for the boosted/high p_T regime. Leptons are either electrons or muons. A veto is imposed on E_T^{miss} in some channels to guarantee orthogonality between final-state channels. The symbol "–" signifies that no requirement is imposed on a given signature. Taken from Ref. [19].

Channel	Diboson state	Selection				VBF cat.	Reference
		Leptons	E_T^{miss}	Jets	b-tags		
$qqqq$	$WW/WZ/ZZ$	0	veto	2J	–	–	[92]
$\nu\nu qq$	WZ/ZZ	0	yes	1J	–	yes	[112]
$\ell\nu qq$	WW/WZ	$1e, 1\mu$	yes	2j, 1J	–	yes	[109]
$\ell\ell qq$	WZ/ZZ	$2e, 2\mu$	–	2j, 1J	–	yes	[112]
$\ell\ell\nu\nu$	ZZ	$2e, 2\mu$	yes	–	0	yes	[97]
$\ell\nu\ell\nu$	WW	$1e$+1μ	yes	–	0	yes	[94]
$\ell\nu\ell\ell$	WZ	$3e$, $2e$+1μ, $1e$+2μ, 3μ	yes	–	0	yes	[108]
$\ell\ell\ell\ell$	ZZ	$4e$, $2e$+2μ, 4μ	–	–	–	yes	[97]
$qqbb$	WH/ZH	0	veto	2J	1, 2	–	[96]
$\nu\nu bb$	ZH	0	yes	2j, 1J	1, 2	–	[93]
$\ell\nu bb$	WH	$1e, 1\mu$	yes	2j, 1J	1, 2	–	[93]
$\ell\ell bb$	ZH	$2e, 2\mu$	veto	2j, 1J	1, 2	–	[93]
$\ell\nu$	–	$1e, 1\mu$	yes	–	–	–	[88]
$\ell\ell$	–	$2e, 2\mu$	–	–	–	–	[101]

The mass distributions and associated systematic uncertainties from the channels considered were combined taking correlations into account. For each signal model tested, only the channels sensitive to that hypothesis were included in the combination as summarized in Table 3.4. Results were calculated in two different signal parameterization paradigms. The first paradigm corresponds to usual one-dimensional upper limits on the cross section times branching fraction as shown in Figure 3.5 for HVT and 3.6 for the bulk RS and scalar models. Each of the channels presented here contributes uniquely to the search for heavy resonances, and the results obtained by their combination extend the reach beyond that of the individual search presented in the previous section.

TABLE 3.4: Signal models, resonances, and decay modes considered in the combination. Taken from Ref. [19].

Model \ Decay mode	WW	WZ	ZZ	WH	ZH	$\ell\nu$	$\ell\ell$
HVT	Z'	W'		W'	Z'	W'	Z'
Bulk RS	G_{RS}		G_{RS}				
Scalar	Scalar		Scalar				

The second paradigm consisted in two-dimensional limits on the coupling strengths of heavy resonances to SM particles in the context of the HVT model, introduced in Section 3.1. These constrains are shown in Figure 3.7 for resonances of mass 2, 3, 4, or 5 TeV. The VV/VH constrains are better at large couplings for g_f and g_H, but become weak as the couplings

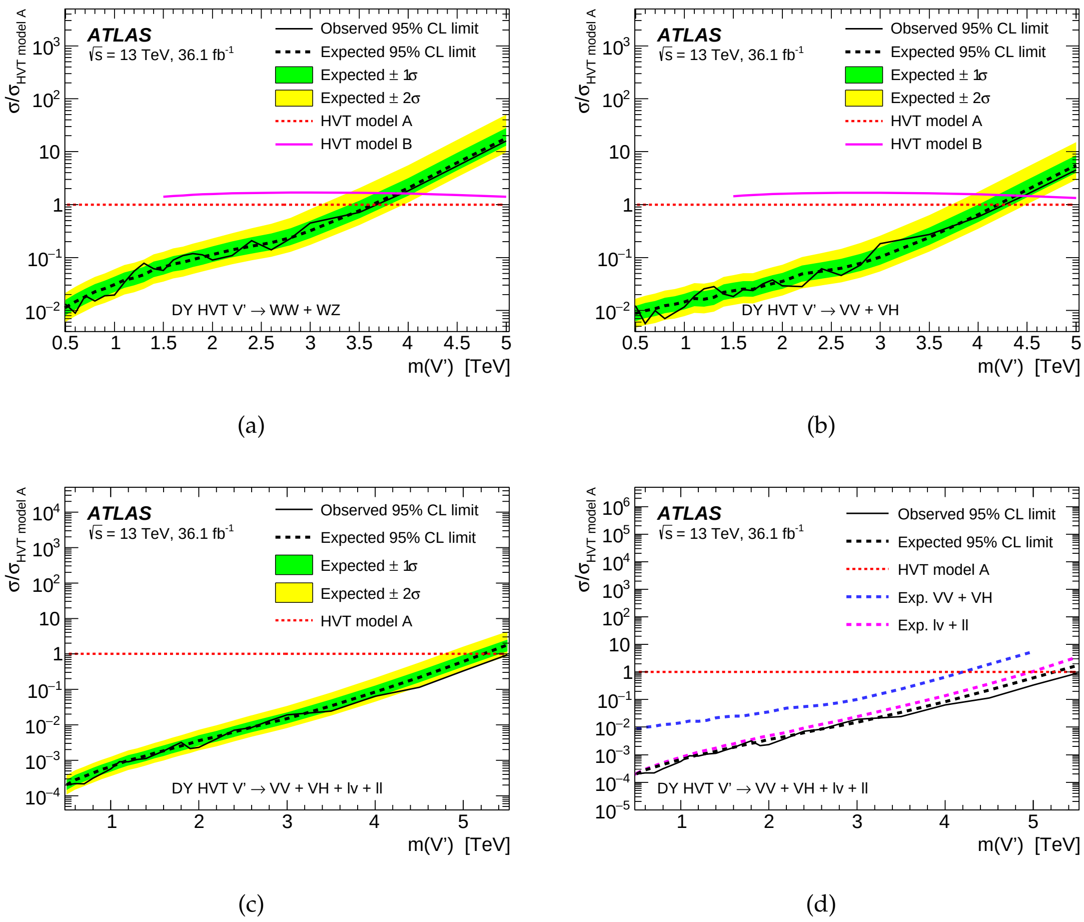


FIGURE 3.5: Observed and expected 95% CL upper limits on the V' cross section times branching fraction for the HVT benchmark model, relative to the cross section times branching fraction for DY production HVT models A or C to: (a) WW or WZ, (b) VV + VH, (c) VV + VH + lepton-antilepton and (d) shows expected limits for bosonic and leptonic decay modes separately. The model predictions are also shown [19].

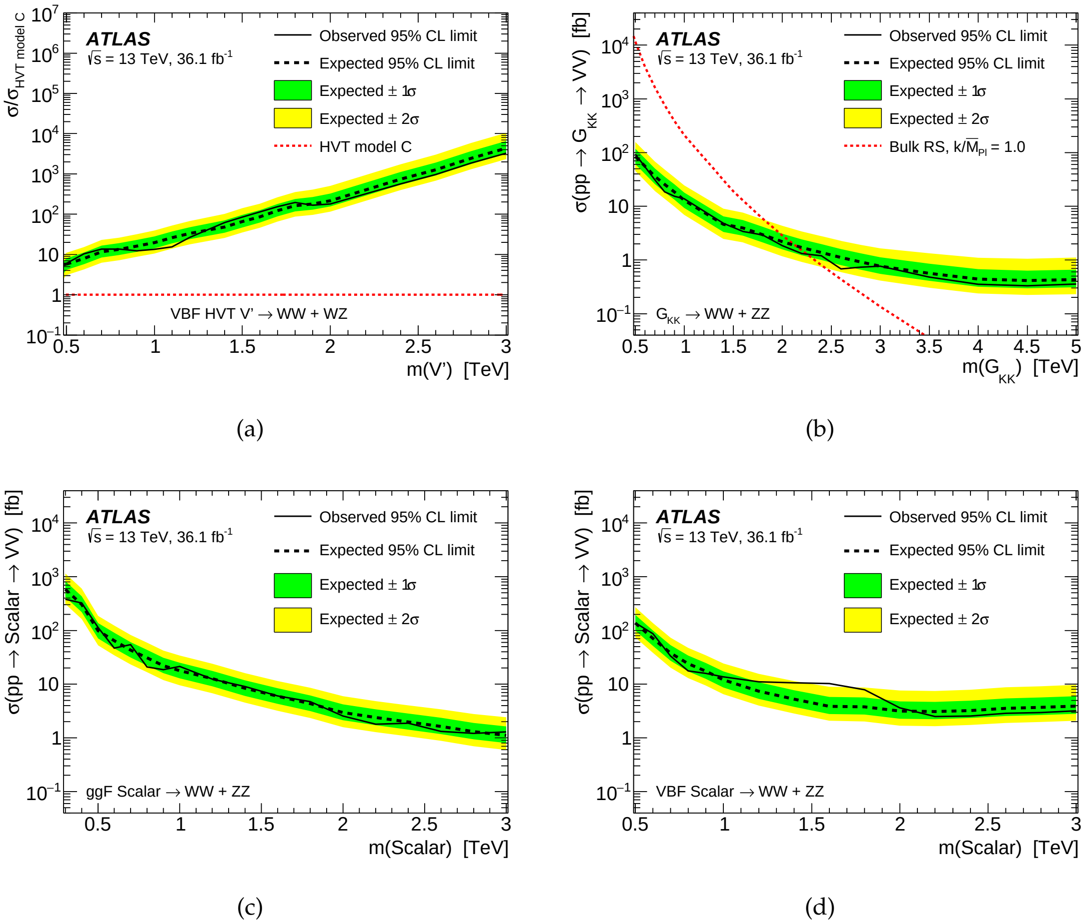


FIGURE 3.6: Observed and expected 95% CL upper limits on the cross section times branching fraction for: (a) the HVT benchmark model VBF production mechanismto to WW or WZ, (b) G_{RS} bulk RS model with $k/\bar{M}_{\rm Pl} = 1$ to WW or ZZ (c) scalar to WW or ZZ for DY production and (d) scalar to WW or ZZ for VBF production. The model predictions are also shown [19].

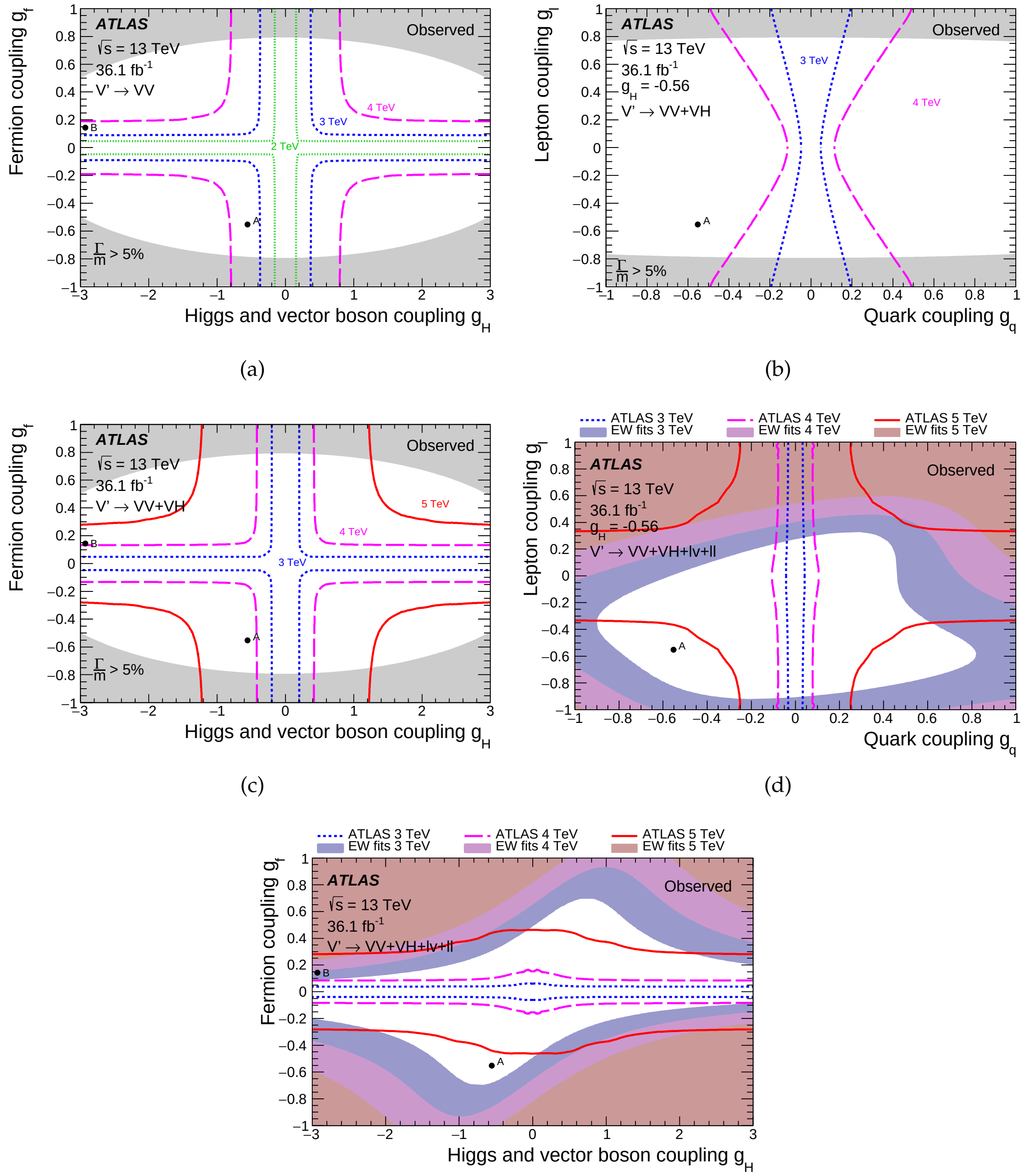
approach zero. The reason is that the resonance couplings to VV and VH tend to zero as the g_H coupling approaches zero, and production of the resonance also tends to zero as the g_f coupling approaches zero. In the case of the $\{g_q, g_\ell\}$ in Fig. 3.7(b) the constraint is worst at large g_ℓ as the bosonic branching fraction decrease while the leptonic one increases in that region. The complementarity is further evidenced by the full $VV/VH/\ell\nu/\ell\ell$ combination in both the $\{g_H, g_f\}$ plane, as shown in Figure 3.7(a), and the $\{g_q, g_\ell\}$ plane, as shown in Figure 3.7(b). The resulting constraints were very stringent, improving on the limits from the best precision EW measurements available at the time in almost all areas of the respective planes, except at low values of quark couplings where resonance production via quark-antiquark annihilation is suppressed at the LHC.

3.4 Summary

This chapter, as well as the start of the LHC Run-2 data taking, started with an exciting question: could a 2 TeV W' or Z' boson exist? Expectation was high for the analysis team working on the fully hadronic VV resonance search.The beginning of Run-2, with the substantial increase in the LHC's collisions energy from 8 TeV to 13 TeV, meant an unprecedented opportunity for the BSM searches as heavy particles would be made much more easily and the statistics would be doubling every few months.

Unfortunately, by the beginning of 2016 there was no sign of the 2 TeV peak. It was a statistical fluctuation. The combination of different diboson searches using the 2015 collected dataset [111] as well as the later combination using the dataset collected between 2015 and 2016 [19] at $\sqrt{s} = 13$ TeV, both with a higher sensitivity to a 2 TeV W' boson than the Run-1 result, showed a good agreement with the SM background expectation. Personally, it was time to explore new ways to look for new physics and achieve a better understanding of the electroweak symmetry breaking mechanism: the indirect way. More details in the next Chapter.

The diboson topologies continue to be a great candle when looking for new physics through heavy resonance searches. The most recent results of the search for diboson resonances in hadronic final states used the full Run-2 dataset collected between 2015 and 2018, 139 fb^{-1} at $\sqrt{s} = 13$ TeV [90]. The analysis used the novel TCC jets to identify the vector bosons (see Section 1.4.1). The upper limits on the production cross-section times branching ratio to diboson final states for the HVT A (B) model excluded the range 1.3 TeV-3.5 TeV (1.3 TeV-3.8 TeV). The bulk RS model a spin-2 Kaluza-Klein graviton was excluded in the range between 1.3 and 1.8 TeV. These limits were further improved by the updated diboson (VV and VH) and di-lepton ($\ell\nu$ and $\ell\ell$) resonance searches combination using the full Run-2 dataset available in May 2022 [20]. In particular, the 95% CL limit on the mass of V' HVT A (B) model was set to 5.8 TeV (4.5 TeV).



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

FIGURE 3.7: Observed 95% CL exclusion contours in the HVT parameter space (a, c, e) $\{g_H, g_f\}$ and (b, d) $\{g_q, g_\ell\}$ with with g_H set to the value from HVT model A for resonances of mass 3, 4, and 5 TeV for the combination of VV, $VV + VH$, and $VV + VH + \ell\nu/\ell\ell$ channels. The areas outside the curves are excluded, as are the filled regions which show the constraints from precision EW measurements. Also shown are the parameters for models A and B, where applicable. The shaded gray area represets a region where the limits are not valid because resonances would have a width greater than 5% of their mass and therefore the assumed narrow-width approximation breaks down [19].

Chapter 4

Probing the SM with diboson topologies including the Higgs boson

The discovery of the Higgs boson in 2012 through the $\gamma\gamma$, WW and $ZZ \rightarrow 4\ell$ decay channels is one of the most important achievements in the particle physics field and from the LHC physics program to date [74, 153]. This discovery marked the beginning of a precision measurements era as well as searches for all the possible production and decay channels to test the hypothesis of a SM-like Higgs.

The Higgs boson in its dominant decay mode, $H \rightarrow b\bar{b}$, was particularly elusive and its observation happened 6 years later in 2018 [75, 154]. This was only possible by reducing the overwhelming background from the strong-interaction production of b-quark pairs by requiring the Higgs boson to be produced in association with a leptonically-decaying W or Z boson, $VH, H \rightarrow b\bar{b}$ production. Given my interest in using diboson topologies for new physics searches I got involved in this analysis in 2019. This time the search for new physics was done in an indirect way, searching for deviations from SM predictions that would point to new physics processes in the Higgs boson production rates.

New phenomena might exist at energy scales beyond LHC's energy reach. Looking for new particles with masses too heavy to be directly produced at LHC is not possible but their existence might be inferred by looking at deviations in different physics processes. In particular, in the case of the Higgs boson, the larger its transverse momentum is, the greater we believe is the sensitivity to new physics processes from new heavy particles. Without knowing the exact nature of this new physics, the LHC experiments have adopted an approach that allow to analyse their data to constrain new types of interactions in the framework of an effective field theory (EFT). The effects can be subtle, in particular if the new physics phenomena are far beyond LHC's energy reach. Therefore it is important to have as many physics analysis inputs as possible to constraint the EFT and $H \rightarrow b\bar{b}$ is one of them.

Capitalizing on my work on boosted $H \rightarrow b\bar{b}$ described in Section 1.4.5 and on $VH, H \rightarrow b\bar{b}$ resonances searches briefly mentioned in Section 3.3 then I joined a new version of the analysis designed to probe the boosted SM Higgs boson, $H \rightarrow b\bar{b}$ production. The results of the first specific boosted $H \rightarrow b\bar{b}$ boson production in association with a leptonically-decaying W or Z boson were subject of a publication [58] in 2020, for which I was one of the paper editors. From October 2019 to October 2021 I was one of the coordinators of the Flavour Tagging Calibration subgroup. I was responsible for the coordination and supervision of all results of this subgroup related to the calibration of the b-tagging and $X \rightarrow b\bar{b}$ algorithms used by the collaboration. Although involved through my role in coordinating and supervising students in many analyses, my personal work during this period focused primarily on the calibration of boosted $H \rightarrow b\bar{b}$ taggers [30] described in Section 4.3.1. Two physics processes are used for the signal calibration of this tagger:

- $Z(\rightarrow b\bar{b}) + \gamma$. An early exploration study to study the potential of this channel for the signal calibration was done with Miles Wu, at the time PhD at the University of Chicago [41]. I co-supervised also his work on the measurement of the jet mass in high transverse momentum $Z(b\bar{b})\gamma$ production [61].

- $Z(\rightarrow b\bar{b})$+jets [30], a work done with Yajun He, PhD student from LPNHE, whose work I am co-directing since October 2019. Yajun He is currently involved in the next edition of the boosted $VH, H \rightarrow b\bar{b}$ using for the first time fully hadronic topologies.

This chapter presents a very brief introduction to the SM Higgs boson in Section 4.1 and an overview of the boosted VH,$H \rightarrow b\bar{b}$ analysis in Section 4.2. Section 4.3 summarises a few selected results relevant for the future version of the boosted $H \rightarrow b\bar{b}$ analyses related with the calibration of the most recent boosted $H \rightarrow b\bar{b}$ tagger. Finally Section 4.4 briefly summarizes the current context of Higgs studies and combinations in the ATLAS experiment. The text in this Chapter is largely based on the References mentioned above.

4.1 The Standard Model Higgs boson

Before 2012, the last particle to complete the SM list of particles was the Higgs boson. This boson is the consequence of introducing a new doublet of complex scalar fields in the SM theory in order to give mass to the electroweak bosons V ($V = W, Z$) and fermions, known as spontaneous electroweak symmetry breaking via the Brout-Englert-Higgs (BEH) mechanism [185, 212, 217].

The scalar sector is represented by the Higgs boson and the scalar potential. The potential has quadratic and quartic terms related to the Higgs boson mass m_H and self-interaction, with coefficients μ^2 and λ, respectively, with $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v$. The couplings between the Higgs boson and the electroweak bosons, g_V, are given by: $g_V = 2\frac{m_V^2}{v}$, where m_V is the electroweak boson mass. The couplings between fermions and the Higgs boson (g_f) are free parameters of the theory but are related to the fermion masses m_f via $g_f = \sqrt{2}\frac{m_f}{v}$.

4.1.1 Higgs boson production at the LHC

The main single Higgs production processes at the LHC are the following: gluon fusion (ggF or $pp \rightarrow H$ is the dominant mode, accounting for almost 90% of the total cross section), vector-boson fusion (VBF or $pp \rightarrow qqH$ is the subleading mode), associated production with a vector boson ($pp \rightarrow VH$), associated production with a top or bottom-quark pair ($pp \rightarrow ttH, bbH$) and associated production with a single top quark ($pp \rightarrow tH$). To date the ggF, VBF, VH and ttH single Higgs production have been observed at the LHC.

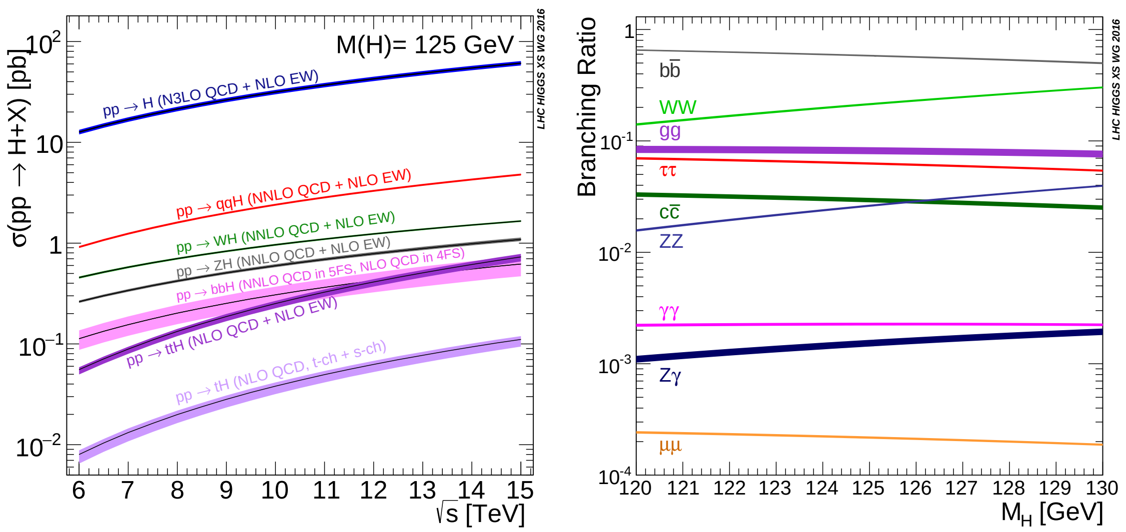
The double Higgs boson production plays also an important role in the understanding of the BEH mechanism. The ggF and VBF double Higgs production are tightly related with the doublet structure of the scalar field and the shape of the potential. For the the production of Higgs pairs the dominant process is gluon fusion. The contribution of a top-quark loop interferes destructively with the diagram that involves the Higgs boson self-coupling. The resulting cross section is more than three orders of magnitude smaller than that of a single Higgs boson, making its measurement a long-term effort. Current projections estimate that a potential evidence of this process at the end of the Run 3 (20220-2025) of the LHC could be at hand. Figure 4.1 shows the cross-section single Higgs production processes as a function of the center-of-mass energy.

Our main interest for this Chapter is the associated production with a W or Z boson ($pp \rightarrow VH$, $V = W, Z$). Most of the time this production mode is accessed with leptonically decaying vector bosons (including neutrinos) which are easier to trigger and quarks to a lesser extent. The quark-initiated processes are calculated up to NNLO QCD and NLO EW accuracies [130, 132, 140, 147, 179–181], with uncertainties below one percent, with additionally 2% uncertainty due to PDFs and α_S. Gluon-initiated processes are sizeable for ZH final states ($\sim$ 10%) and bring the total uncertainties to about 4% and they are calculated up to NLO+NLL [6, 131, 189–191].

4.1.2 Higgs boson decay modes

The branching ratios (BR) of the main decay modes of the Higgs boson are shown in Figure 4.1. The largest BR corresponds to the $b\bar{b}$ decay. However due to the very large multi-jet background at the LHC, decay modes with hadronic final states are difficult to access and therefore so far this decay has been observed in the VH production mode, with W and Z bosons decaying leptonically [18, 58, 68] but studied in several production modes including ggF, VBF and ttH. Clean signatures like $\gamma\gamma$, $ZZ \to 4\ell$ ($\ell = e, \mu$) and $WW \to \ell\nu\ell\nu$ have higher sensitivities despite the smaller branching ratio and can assess all the production modes. These were the three discovery channels and now we are doing precision measurements with them.

The $\tau\tau$ decay is studied predominantly in the VBF production or in association with a high-p_{T} jet from ggF. It was was observed with the 2015-2016 data [26, 66, 155] providing new insight into the Yukawa coupling. Similarly to the $b\bar{b}$ decay, the cc decay give physicists new insight into the Higgs boson's relationship with the second generation of matter particles and is also heavily dominated by background processes. A recent search study for $VH, H \to c\bar{c}$ was published showing no significant sign of the decay of the Higgs boson to charm quarks yet but improved constraints on the charm Yukawa coupling modifier [28, 149]. Finally, decays to $\mu\mu$ and $Z\gamma$ will require large luminosities to be observed.



(a) (b)

FIGURE 4.1: (a) Higgs boson production cross sections as a function of the centre-of-mass energy and (b) Higgs boson decay branching ratios as a function of its mass [200].

4.1.3 More about the beauty Higgs

$H \to b\bar{b}$ is an ideal decay channel to search for deviations in the production rate from predictions that would point to new physics processes. As the most likely decay of the Higgs boson (accounting for $\approx 58\%$ of all Higgs-boson decays), its larger abundance allows physicists to probe further into the high-transverse-momentum regions, where the production rate decreases due to the composite structure of the colliding protons. This channel was first observed in 2018 [75, 154] and provided direct evidence for the Yukawa coupling of the Higgs boson to down-type quarks using pp collisions of the Run-1 data set combined with a partial Run 2 data set collected at 13 TeV by both experiments. A first cross-section measurement as a function of the vector-boson transverse momentum in the $VH, H \to b\bar{b}$ was also carried out by the ATLAS Collaboration using 79.8 fb^{-1} of Run-2 data [64].

Since then the full Run-2 dataset was analysed by two different analysis made public in 2020. One targeting events with vector bosons with transverse momentum $p_{\mathrm{T}}^{V} > 75$ GeV which reconstructs the Higgs boson candidate from two jets with radius parameter of $R = 0.4$, referred to as small-radius (small-R) jets [68]. And a second one, targeting specifically (and for the first time) the high transverse momentum regime $p_{\mathrm{T}}^{V} > 250$ GeV [58]. This high-p_{T} regime is particularly interesting due to its sensitivity to beyond the Standard Model anomalous couplings in models of effective field theories [133, 168, 175, 203]. For boosted Higgs, the decay products can become close enough that they cannot be reconstructed with two small-R jets. To explore this 'boosted' regime, the Higgs boson is reconstructed as a single large-R jet with $R = 1.0$ [134]. These two full Run-2 $VH, H \to b\bar{b}$ measurements were combined recently in 2021 [18], providing improved precision at moderate p_{T}^{V} between 75 and 250 GeV, and adding differential information for $p_{\mathrm{T}}^{V} > 250$ GeV.

The $H \to b\bar{b}$ process has been also explored in other production modes in ATLAS and CMS: gluon fusion boosted production [25, 151], VBF [67, 98] and ttH [56, 156] (for which an evidence was obtained by CMS with 3.9σ).

4.2 The boosted VH, $H \to b\bar{b}$ analysis at 13 TeV

Three analysis channels based on the vector boson decay were defined: 0-lepton targeting $ZH \to \nu\nu b\bar{b}$, 1-lepton targeting $WH \to \ell\nu b\bar{b}$ and 2-lepton targeting $ZH \to \ell\ell b\bar{b}$ (with ℓ= electron or muon). The Higgs boson is reconstructed as a single large-R jet and the b-quarks from its decay as a pair of jets, reconstructed with a p_{T}-dependent radius parameter, associated with the large-R jet and identified as containing a b-hadron.

The dominant background processes after the event selection correspond to: $Z +$ jets and $t\bar{t}$ events with a significant contribution from $W +$ jets and diboson production in the 0-lepton channel; $t\bar{t}$ and $W +$ jets production followed by the single-top background in the 1-lepton channel and $Z +$ jets production is the dominant background followed by the ZZ background in the 2-lepton channel. The signal is extracted from a combined profile likelihood fit to the large-R jet mass, using several signal and control regions. The yield of diboson production VZ with $Z \to b\bar{b}$ is also measured using the same fit and provides a validation of the analysis. The cross-section measurements are performed within the simplified template cross-section (STXS) framework [8, 200]. The STXS categories are defined by the production process, WH or ZH, where the latter contains contributions from both $qq \to ZH$ and $gg \to ZH$ diagrams and they are defined for different p_{T}^{V} bins. These measurements are then used to constrain anomalous couplings from beyond SM scenarios using the EFT framework as a benchmark.

4.2.1 Monte Carlo simulation and data

The analysis was performed using the dataset collected in 2015-2018 from 13 TeV LHC pp collisions using a combination of triggers selecting events with high missing transverse momentum or with a high-p_{T} lepton, depending on the analysis channel. More details can be found in Ref. [58]. The integrated luminosity of this dataset after requiring good beam and detector conditions is 139 fb^{-1}, with a relative uncertainty of $\pm 1.7\%$ [54], obtained using the LUCID-2 detector [55, 124] for the primary luminosity measurements.

MC simulated event samples are used to model the signal and background contributions, except for the multijet production, whose contribution is estimated with data-driven techniques. The following signal processes were generated: $qq \to WH \to \ell\nu b\bar{b}$, $qq \to ZH \to \nu\nu b\bar{b}/\ell\ell b\bar{b}$ and $gg \to ZH \to \nu\nu b\bar{b}/\ell\ell b\bar{b}$. The signal and top quark processes were generated with the POWHEG-BOX v2 [3] generator interfaced to PYTHIA 8.212 [213] for the parton shower simulation. V+jets and diboson processes were generated with the SHERPA 2.2.1 [141, 186, 187] generator, except for gluon-initiated diboson production where version 2.2.2 is used.

The samples are normalised according to the predictions described in Section 4.1. Full details on the nominal generators and parameters used as well as the samples produced with alternative generators used to estimate modelling systematic uncertainties are given in Ref. [58]. All samples are passed through a full detector simulation [117] based on GEANT4 [2], including pile-up from additional pp collisions in the same or neighbouring bunch crossings generated with PYTHIA 8.1 [213]. Simulated events were then reconstructed with the same algorithms as those applied to data and are weighted to match the pile-up distribution observed in the data.

4.2.2 Object and event selection

The analysis relies on the reconstruction of electrons, muons, jets, and missing transverse momentum (with magnitude $E_{\text{T}}^{\text{miss}}$). The small differences between the efficiencies measured in data and MC simulation for these objects are corrected for by applying scale factors to the MC simulation so that it matches the data.

Events are categorised into the 0-, 1- and 2-lepton channels depending on the number of selected electrons and muons to target the $ZH \to \nu\nu b\bar{b}$, $WH \to \ell\nu b\bar{b}$ and $ZH \to \ell\ell b\bar{b}$ signatures, respectively. *Baseline* electrons are required to have $p_{\text{T}} > 7$ GeV and $|\eta| < 2.47$, to be isolated from other tracks and energy deposit clusters, to meet loose likelihood selection criteria based on shower shapes and $|d_0/\sigma(d_0)| < 5$ and $|z_0 \sin\theta| < 0.5$ mm. *Signal* electrons are a subset of the baseline electron set and are selected using a tighter likelihood requirement and and using a tighter calorimeter-based isolation criterion [32] . Muons are required to have $p_{\text{T}} > 7$ GeV and $|\eta| < 2.7$ and to have $|d_0/\sigma(d_0)| < 3$ and $|z_0 \sin\theta| < 0.5$ mm. Two muon categories are used in the analysis: *baseline* muons are selected using the 'loose' identification criterion and a loose track isolation; *signal* muons are required to have $|\eta| < 2.5$, to satisfy the 'medium' identification criterion and a tighter track-based isolation criterion [72]
.

Three jets collections are used in the analysis:

- Small-R jets, anti-k_t $R = 0.4$ jets build from topoclusters, are used in building the missing transverse momentum and event categorisation.

- Large-R jets, anti-k_t $R = 1.0$ trimmed jets with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$ also built from topocluster at the LC scale, are used to reconstruct the boosted Higgs candidate. Large-R jets are required to have $p_{\text{T}} > 250$ GeV, combined mass $m_{\text{J}} > 50$ GeV and $|\eta| < 2.0$, the last due to tracking acceptance.

- Track-jets, anti-k_t variable-radius jets, are used to reconstruct a candidate two-body $H \to b\bar{b}$ decay within the large-R jet. Only VR track-jets with $p_{\text{T}} > 10$ GeV, $|\eta| < 2.5$ and with at least two constituents are considered [65]. VR track-jets are matched to the

TABLE 4.1: Event selection requirements for the boosted $VH, H \to b\bar{b}$ analysis channels and sub-channels [58].

Selection	0 lepton channel	1 lepton channel		2 leptons channel	
		e sub-channel	μ sub-channel	e sub-channel	μ sub-channel
Trigger	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single electron	$E_{\mathrm{T}}^{\mathrm{miss}}$	Single electron	$E_{\mathrm{T}}^{\mathrm{miss}}$
Leptons	0 *baseline* leptons	1 *signal* lepton $p_{\mathrm{T}} > 27$ GeV no second *baseline* lepton	$p_{\mathrm{T}} > 25$ GeV	2 *baseline* leptons among which ≥ 1 *signal* lepton, $p_{\mathrm{T}} > 27$ GeV both leptons of the same flavour -	opposite sign muons
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 250 GeV	> 50 GeV	-	-	
p_{T}^{V}	$p_{\mathrm{T}}^{V} > 250$ GeV				
Large-R jets	at least one large-R jet, $p_{\mathrm{T}} > 250$ GeV, $\lvert\eta\rvert < 2.0$				
Track-jets	at least two track-jets, $p_{\mathrm{T}} > 10$ GeV, $\lvert\eta\rvert < 2.5$, matched to the leading large-R jet				
b-tagged jets	leading two track-jets matched to the leading large-R must be b-tagged (MV2c10, 70%)				
m_{J}	> 50 GeV				
$\min[\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}$, small-$R$ jets$)]$	$> 30°$	-			
$\Delta\phi(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, H_{\mathrm{cand}})$	$> 120°$	-			
$\Delta\phi\,(\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}, \boldsymbol{E}_{\mathrm{T,\,trk}}^{\mathrm{miss}})$	$< 90°$	-			
$\Delta y(V, H_{\mathrm{cand}})$	-	$\lvert\Delta y(V, H_{\mathrm{cand}})\rvert < 1.4$			
$m_{\ell\ell}$	-			66 GeV$< m_{\ell\ell} < 116$ GeV	
Lepton p_{T} imbalance	-			$(p_{\mathrm{T}}^{\ell_1} - p_{\mathrm{T}}^{\ell_2})/p_{\mathrm{T}}^{Z} < 0.8$	

large-R calorimeter jets via ghost-association [1]. Track-jets not associated with large-R jets are also used in the analysis for event categorisation as will be mentioned later. To identify track-jets containing b-hadron decay products, track-jets are tagged using the MV2c10 algorithm introduced in Section 1.4.5, using the 70% efficiency working point.

The energy and mass of the large-R and small-R jets are then calibrated following the procedure described in Section 1.4.4. Two additional corrections are applied to the large-R jets in order to improve the scale and the resolution of their energy and mass measurements. The first one corrects for the energy of the muons from semileptonic decays of the b-hadrons represented by b-tagged track-jets. For the second one a per-event likelihood uses the full reconstruction of the event kinematics to improve the estimate of the energy of the b-jets [34] and it is only applied in the 2-lepton channel.

To select the Higgs boson candidate, the leading p_{T} large-R jet is chosen, at least two VR track-jets are required to be matched to it by ghost-association, and the two leading VR track-jets are required to be b-tagged. More details about this identification method can be found in Section 1.4.5. The large-R jet's invariant mass (m_{J}) is used as the discriminant. The reconstructed transverse momentum p_{T}^{V} of the vector boson corresponds to $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 0-lepton channel (as $E_{\mathrm{T}}^{\mathrm{miss}}$ is a proxy of the undetected neutrinos from the Z boson decay), to the magnitude of the vector sum of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the muon or electron in the 1-lepton channel, and to the transverse momentum of the two muons or two electrons system in the 2-lepton channel. The full event selection is detailed in Table 4.1. The additional cuts in the 0-, 1- and 2-lepton channels were optimised to remove multijet background from jet energy mismeasurements, non-prompt lepton background and the main Z+jets background, respectively.

The selected events in the 3 lepton categories are categorised in ten signal regions (SRs) and the four control regions (CRs) enriched in background events. The categorization is summarised in Table 4.2. Events are split into two p_{T}^{V} bins with $250 < p_{\mathrm{T}}^{V} < 400\,$GeV and with $p_{\mathrm{T}}^{V} \geq 400\,$GeV since the signal-to-background ratio increases for large Higgs boson transverse momenta [16, 134]. Also, the 250 GeV and 400 GeV thresholds align with the fiducial definition of the VH STXS categories. The selection efficiency in the 0-, 1- and 2-lepton channels and two p_{T}^{V} bins ranges between approximately 6% and 16% for the WH and ZH processes where the W and Z bosons decay leptonically and the Higgs boson decays into a pair of b-quarks. CRs enriched in $t\bar{t}$ are built in the 0- and 1-lepton channels if there is at least one additional b-tagged track jet, which is not associated to the Higgs candidate jet. Events in the 0- and 1-lepton SRs are further categorised depending on the number of small-R jets not matched to the Higgs-jet candidate, i.e. with $\Delta R(H_{\mathrm{cand}}$, small-$R$ jet)>1.0. If no additional jets are present, the event falls into the *high-purity* signal region (HP SR), and if at least one additional jet is present into the *low-purity* signal region (LP SR).

TABLE 4.2: Summary of the definition of the analysis regions. Signal enriched regions are marked with the label SR. There are regions with relatively large signal purity (HP SR) and with low purity (LP SR). Background enriched regions are marked with the label CR. The shorthand "add" stands for additional small-R jets, i.e. number of small-R jets not matched to the Higgs-jet candidate [58].

Channel	Categories					
	$250 < p_{\mathrm{T}}^{V} < 400$ GeV			$p_{\mathrm{T}}^{V} \geq 400$ GeV		
	0 add. b-track-jets		≥ 1 add. b-track-jets	0 add. b-track-jets		≥ 1 add. b-track-jets
	0 add. small-R jets	≥ 1 add. small-R jets		0 add. small-R jets	≥ 1 add. small-R jets	
0-lepton	HP SR	LP SR	CR	HP SR	LP SR	CR
1-lepton	HP SR	LP SR	CR	HP SR	LP SR	CR
2-lepton	SR			SR		

4.2.3 Results

The results are obtained from a binned maximum-profile-likelihood fit to the data of the m_{J} distribution, using all the SRs and CRs defined in Section 4.2.2. The fit is performed using the ROOSTATS framework [204, 216]. The likelihood function is described in more detail in Ref. [34].

The systematic uncertainties included in the fit as nuisance parameters belong to the following categories: experimental, signal modelling, and background modelling. A full description can be found in Ref. [58]. Experimental systematic uncertainties related to trigger event selection, luminosity, lepton reconstruction and identification, missing transverse momentum, b-tagging correction factors and jet calibration were included. The signal modelling uncertainties took into account the choice of PDF, parton shower model and missing higher order QCD and electroweak corrections. The principal background modelling uncertainties considered were the renormalisation and factorisation scale variations; merging scale variations from multi-leg simulations; resummation scale or parton shower uncertainties; PDF uncertainties; and differences with alternative MC generators. The dominant source of systematic uncertainty for the boosted $VH, H \to b\bar{b}$ analysis is experimental and related to the large-R jet calibration, in particular in the m_{J} resolution, amounting to $\approx 10-13\%$. The second two largest source of systematic uncertainty is the background modelling and the limited size of the MC simulation samples, which overall has an impact of $\approx 10\%$ on the result each. Systematic uncertainties in the signal modelling have an impact of $\approx 4\%$, on par with uncertainties related to small-R jets. The impact of systematic uncertainties is almost as important as the total statistical uncertainty.

Figure 4.2 show the m_{J} distributions after fit for all three channels' SRs. The low-purity and high-purity categories in the case of the 0-lepton and 1-lepton channels are merged in Figure 4.2. In all SRs and CRs a good agreement between the data and the prediction is observed. The m_{J} distribution is shown in Figure 4.3(a) summed over all channels and signal regions, weighted by their respective values of the ratio of the fitted Higgs boson signal to background yields and after subtraction of all backgrounds except for the WZ and ZZ diboson processes.

Signal strength and cross-section measurements

The signal strengths for VZ and VH extracted simultaneously through a maximum-likelihood fit to the m_{J} distribution were found to be:

$$\mu^{bb}_{VH} = 0.72^{+0.39}_{-0.36} = 0.72^{+0.29}_{-0.28}(\text{stat.})^{+0.26}_{-0.22}(\text{syst.}).$$

$$\mu^{bb}_{VZ} = 0.91^{+0.29}_{-0.23} = 0.91 \pm 0.15(\text{stat.})^{+0.24}_{-0.17}(\text{syst.}),$$

With all three lepton channels combined, a significance of 5.4 standard deviations was observed for the VZ process, compared to an expectation of 5.7σ. For a Higgs boson mass of 125 GeV, the observed excess with respect to the background-only hypothesis has a significance of 2.1σ, to be compared with an expectation of 2.7σ. Figure 4.3(b) shows the results of a fit with six VH POIs measuring the individual signal strengths in each of the three channels and p_{T}^{V} bins separately.

The cross-sections in the two defined p_{T}^{V} regions, $250\,\text{GeV} < p_{\mathrm{T}}^{V} < 400\,\text{GeV}$ and $p_{\mathrm{T}}^{V} \geq 400\,\text{GeV}$ were also measured in this analysis. The analysis closely followed the strategy used in Ref. [64]. The fit performed had multiple parameters of interest (POIs) corresponding to the cross-sections in the four regions used in the analysis, multiplied by the $H \to b\bar{b}$ and $V \to$ leptons branching fractions. The measured results, together with the SM predictions in the selected STXS regions, are summarised in Figure 4.4. The measurements are in agreement with the SM expectations.

Constraints on effective interactions

The STXS cross-section results were interpreted in the SMEFT approach to set constraints on potential modifications of the strength and tensor structure of the $VH, H \to b\bar{b}$ process. In this SMEFT approach, the SM Lagrangian is extended with higher-dimensional operators that capture the low-energy limit effects of a fundamental ultraviolet theory, without a priori knowledge of this theory [168]

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{d} \frac{1}{\Lambda^{d-4}} \left(\sum_{i} c_i^{(d)} \mathcal{O}_i^{(d)} \right),$$

where $\mathcal{O}_i^{(d)}$ are dimension-d operators and $c_i^{(d)}$ are the corresponding numerical coefficients called Wilson coefficients. All Wilson coefficients are assumed real. In the SM, all Wilson coefficients are zero. The scale of new physics Λ is a free parameter set to 1 TeV. In this analysis, the Warsaw basis [188] of dimension $d = 6$ operators is used, taking into account only the lepton- and baryon-number-conserving ones. The Wilson coefficients are used to parameterise the STXS and the Higgs boson decay rates [69] from leading-order predictions [133] and can be constrained using the STXS measurements presented in Figure 4.4. The parameterisation of the STXS takes into account the linear terms originating from the interference

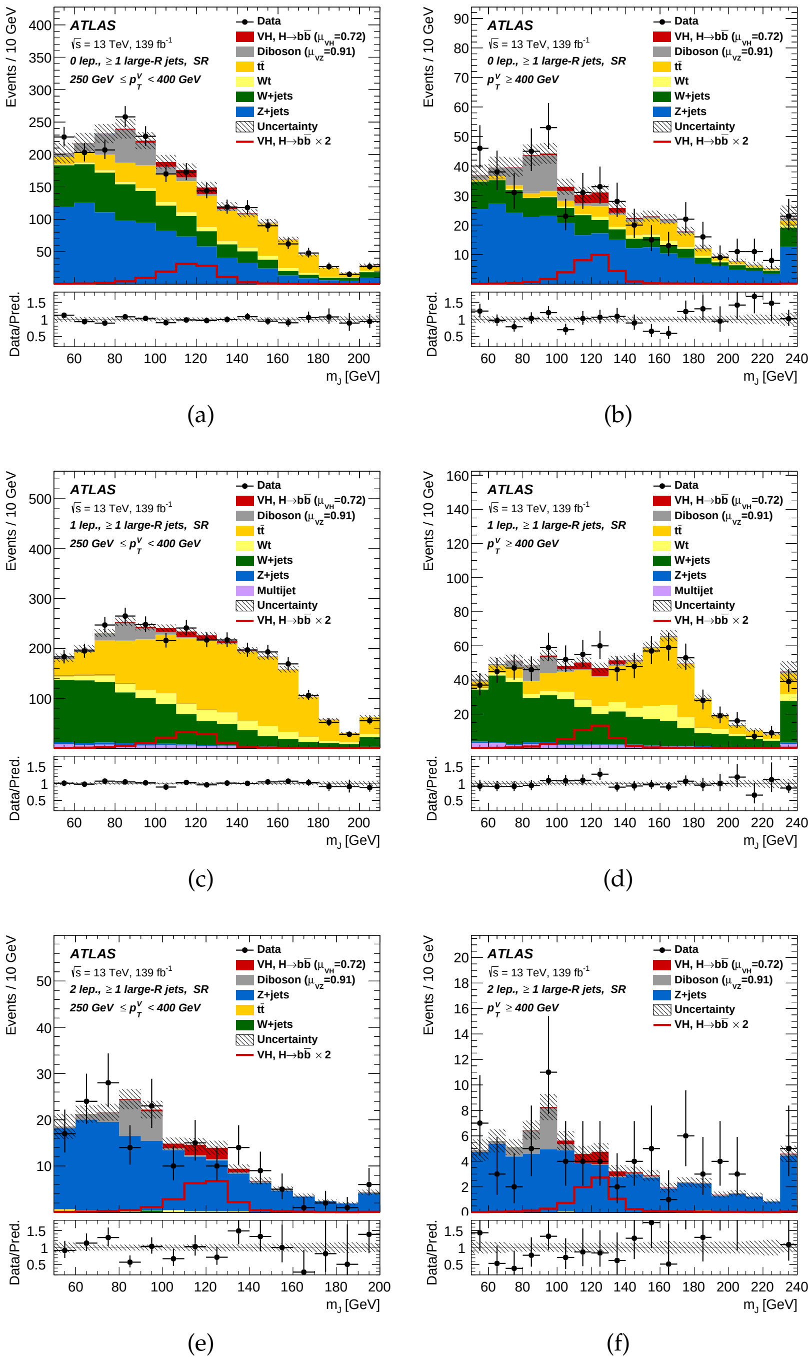
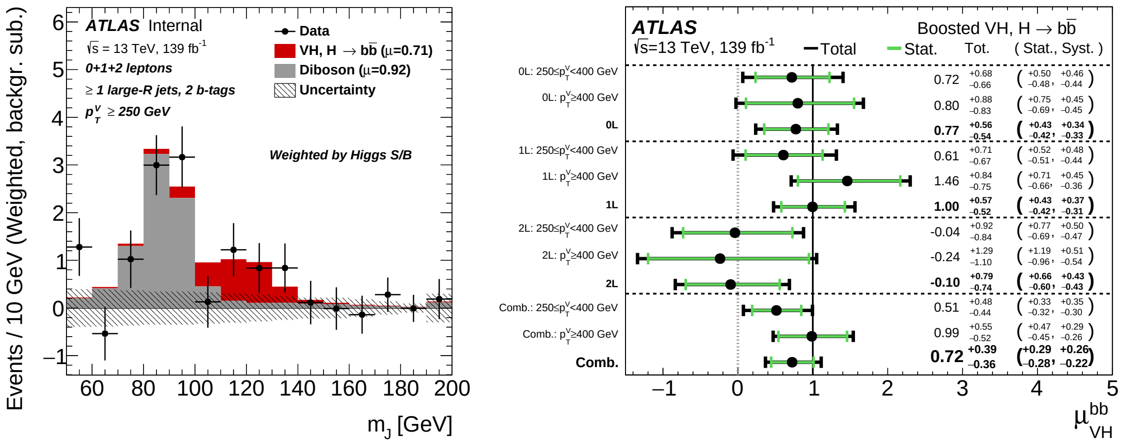


FIGURE 4.2: The m_J post-fit distributions in (a, b) the 0-lepton, (c, d) 1-lepton and (e, f) 2-lepton signal regions for 2-b-tagged events for (a, c, e) $250\,\text{GeV} < p_T^V < 400\,\text{GeV}$ and (b, d, f) $p_T^V \geq 400\,\text{GeV}$. The low-purity and high-purity categories in the case of the 0-lepton and 1-lepton channels are merged in this figure. The background contributions after the likelihood fit are shown as filled histograms. The Higgs boson signal (m_h = 125 GeV) is shown as a filled histogram on top of the fitted backgrounds normalised to the signal yield extracted from data ($\mu_{VH}^{bb} = 0.72$), and unstacked as an unfilled histogram, scaled by the SM prediction times a factor of two. The size of the combined statistical and systematic uncertainty for the sum of the fitted signal and background is indicated by the hatched band. The highest bin in the distributions contains the overflow. The ratio of the data to the sum of the fitted signal and background is shown in the lower panel [58].



(a) (b)

FIGURE 4.3: (a) m_J distribution in data after subtraction of all backgrounds except for the WZ and ZZ diboson processes. The contributions from all lepton channels and signal regions are summed and weighted by their respective values of the ratio of fitted Higgs boson signal and background yields. (b) Fitted values of the signal strength parameter, μ_{VH}^{bb}, for $m_H = 125$ GeV for the 0-, 1- and 2-lepton channels in different p_T^V regions separately and for various combinations [58].

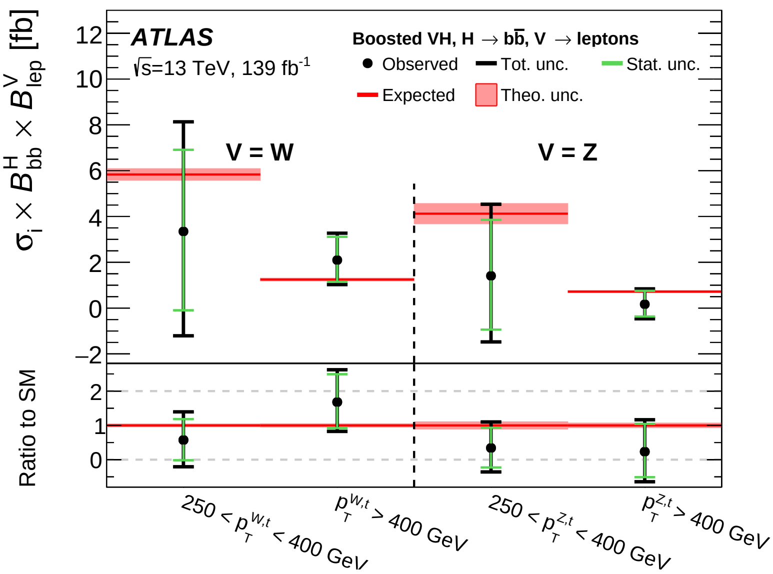


FIGURE 4.4: Measured VH STXS [8, 200] cross-sections times the $H \to b\bar{b}$ and $V \to$ leptons branching fractions for the boosted analysis [58].

between SM and non-SM amplitudes as well as the quadratic ones from the squared non-SM amplitudes.

Figure 4.5 shows the results for the four Wilson coefficients the boosted $VH, H \to b\bar{b}$ analysis is most sensitive to (after removing degeneracies), $c_{Hq}^{(3)}$, c_{Hu}, c_{HW}, and c_{HWB}, in addition to the operator which affects the $H \to b\bar{b}$ decay, $\mathcal{O}_{dH}$. No significant deviation from their expected SM value is observed. More information about the Wilson coefficients exact definition, the negative-log-likelihood one-dimensional profiles for the individual fits as well as other parametrisation approaches for the Wilson coefficients can be found in Ref. **??**.

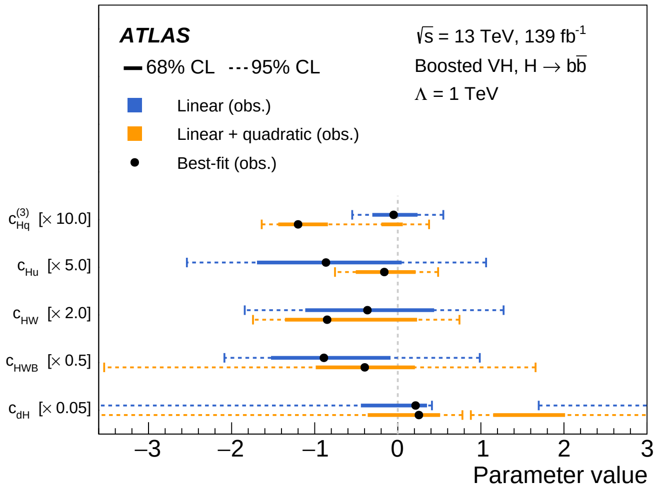
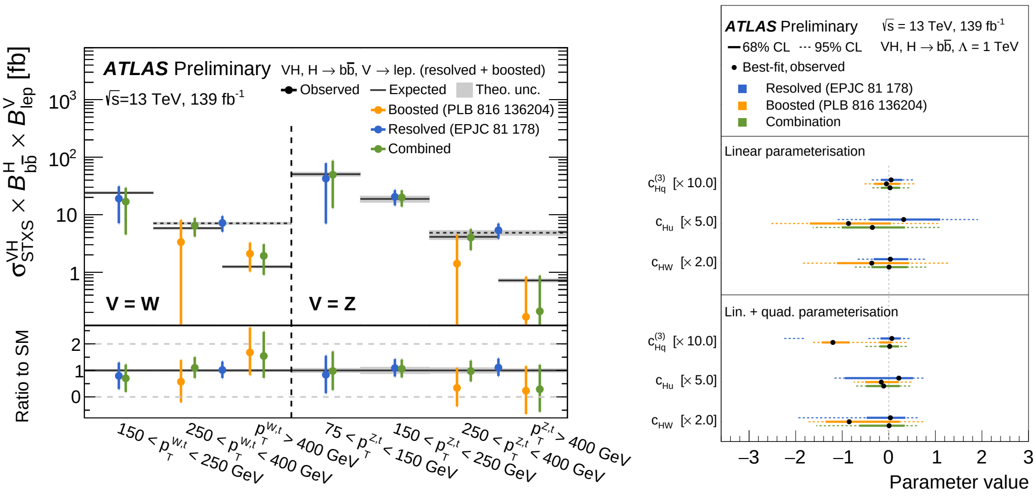


FIGURE 4.5: Summary of the observed individual confidence interval at 68% (solid lines) and 95% (dashed lines) CL for the $c_{Hq}^{(3)}$, c_{Hu}, c_{HW}, c_{HWB} and c_{dH} Wilson coefficients from a fit of the STXS, using a linear-only parameterisation (in blue) and including quadratic terms (in orange) from the boosted $VH, H \to b\bar{b}$ analysis [58].

4.2.4 Combination of the resolved and boosted $VH, H \to b\bar{b}$ analyses

At the same timescale of the publication of the first boosted $VH, H \to b\bar{b}$ analysis, the resolved version was also updated. The new resolved analysis used 75% more data than the previous one matching the full Run-2 dataset also used in the boosted analysis, increased MC simulation statistics and improved event selection. A combination of the two analyses was performed in 2021 [18], after some modifications were introduced to avoid overlapping events.

The resolved analysis provided a more precise measurement of the cross-section in the inclusive $p_{\mathrm{T}}^{V} > 250$ GeV region. This can be attributed to the larger acceptance at lower p_{T}^{V} value, the usage of more precise physics objects calibration and to the use of multivariate analysis techniques. Their combination resulted in a higher precision at moderate p_{T}^{V} between 75 and 250 GeV and added differential information for $p_{\mathrm{T}}^{V} > 250$ GeV. The measured signal strength was $\mu_{VH}^{bb} = 1.00^{+0.18}_{-0.17} = 1.00^{+0.12}_{-0.11}(\text{stat.})^{+0.14}_{-0.13}(\text{syst.})$. Figure 4.6 shows the STXS results for the combination as well as the observed results for the 68% and 95% CL intervals for three representative operators with strong p_{T}^{V} dependence and for which the analyses were sensitive. The combination provides more stringent limits than the individual analyses, demonstrating the benefits of increasing in the number of measured STXS regions for cases where multiple operators are constrained simultaneously.



(a) (b)

FIGURE 4.6: (a) Measured VH STXS cross-sections times the $H \to b\bar{b}$ and $V \to$ leptons branching fractions for the resolved (blue) and boosted (yellow), and the combination (green) analyses. (b) Observed (right) 68% and 95% CL intervals for the three selected Wilson coefficients for the resolved (blue) and boosted (yellow), and the combination (green) $VH, H \to b\bar{b}$ analyses. [18].

CMS released their latest $VH, H \to b\bar{b}$ during the summer 2022 [166] including categories for resolved and boosted Higgs boson decays with optimised overlap removal among them. The measured inclusive signal strength was $\mu_{VH}^{bb} = 0.58^{+0.19}_{-0.18}$ leading to an observed (expected) significance of 3.3σ (5.2σ). The boosted Higgs candidate is tagged using a performant DeepAK8 algorithm [167] leading to a good sensitivity for $p_{\mathrm{T}} > 400$ GeV. DeepAK8 uses low-level input features (PFlow candidates) in addition to the traditional observables of the boosted jet environment (tracks, jets, secondary vertices). Figure 4.7 shows the STXS results for the CMS $VH, H \to b\bar{b}$ analysis using the full Run-2 dataset.

4.3 Preparing the legacy Run-2 boosted $H \to b\bar{b}$ analyses

Further improvements of the analysis presented in Section 4.2 can come from: additional data (Run-3 data taking period just started during the summer of 2022), the addition of new VH channels/topologies (more about this in Chapter 5) and/or improvements in the analysis strategy.

In this section we will focus on one of the possible improvements in the analysis strategy: the use of the new boosted $H \to b\bar{b}$ using neural networks, the so called $X \to b\bar{b}$ tagger, described in Section 1.4.5. When compared to double b-tagging algorithms used by ATLAS in Run-2 (also described in Section 1.4.5), the new NN tagger improves the discrimination against top and multijet backgrounds, specially at high p_{T}. Relative to a double MV2 tag used in Section 4.2, the multijet (top) rejection at 60% efficiency is increased by a factor of 1.4 (2.0) for jets with $p_{\mathrm{T}} > 500$ GeV, see Figure 1.13.

After the taggers are developed and trained based on MC simulation, their validation with ATLAS data is necessary for their application in physics data analysis since their efficiency might differ between data and simulation. The intrinsic abundance of hadronic processes at the LHC make this task particularly challenging. In the latest ATLAS results, the $X \to b\bar{b}$ tagger signal efficiency is calibrated using $Z(\to b\bar{b})\gamma$ and $Z(\to b\bar{b})$+jets. $Z \to b\bar{b}$

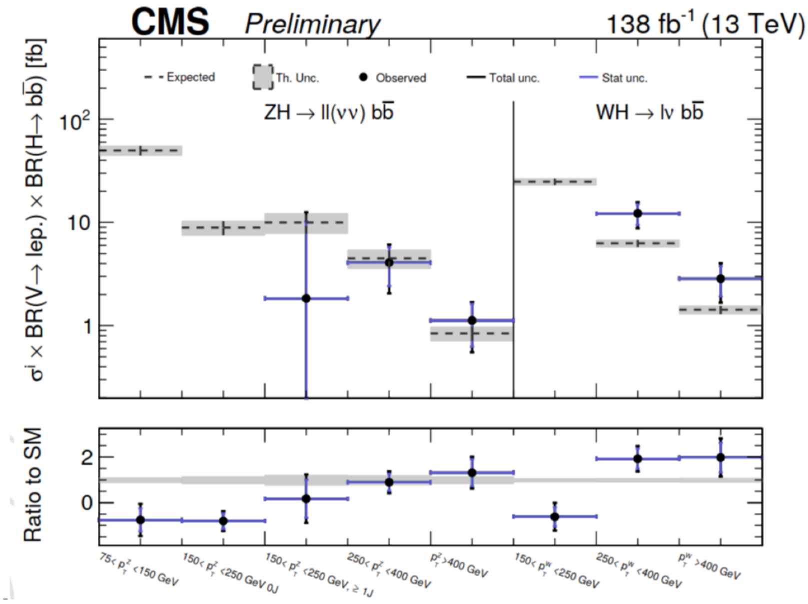


FIGURE 4.7: Measured VH STXS [8, 200] cross-sections times the $H \to b\bar{b}$ and $V \to$ leptons branching fractions by CMS [166].

process is a colour-singlet resonance with a mass close to the Higgs boson mass, so kinematic properties of the $Z \to b\bar{b}$ and $H \to b\bar{b}$ events are expected to be similar. The background efficiency is calibrated using $t\bar{t}$ events. In addition the use of multijet events enriched in $g \to b\bar{b}$ splitting for background mistagging efficiency calibration has been also explored and the modelling of large-R jet kinematics in Monte Carlo simulation is checked after the application of the tagger in those events. These two backgrounds are usually significant for analyses searching for $H \to b\bar{b}$ events in fully hadronic final states. I have actively contributed to the development of the signal efficiency methodology in ATLAS and therefore this section will focus on it. More information about the background mistagging efficiency calibrations can be found in Ref. [30].

Between 2015 and 2017 with Miles Wu, (at the time) PhD student at University of Chicago (2016-2017), we explored the possibility of using $Z \to b\bar{b}$ events for the calibrating the available $H \to b\bar{b}$ two b-tagged track-jet tagger. The calibration based on $Z(\to b\bar{b})\gamma$ events was first considered by ATLAS in our paper Ref. [41] for which I was one of the editors. The $Z\gamma$ studies were pursued later on, leading to the a measurement of the fiducial and differential jet mass cross-sections for this process [61] in a collaboration between the University of Chicago and the University of Arizona. While previous analyses [63, 152, 162] measured the cross-section of quark and gluon-initiated jets for different grooming algorithms, in this analyses we measured for the first time the mass of large-R jets containing the hadronic decay products of Z bosons in $Z(\to b\bar{b})\gamma$ events using 36 fb^{-1} of data collected at $\sqrt{s} = 13$ TeV for two jet collections: anti-k_t $R = 1.0$ jets trimmed with $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.2$ and soft-drop with $\beta = 0$ and $z_{\text{cut}} = 0.1$ SK+CS. The observed experimental features of these different grooming algorithms are helpful for future measurements of resonant $b\bar{b}$ decays but also for comparison with future precision theoretical calculations of the jet mass spectra.

The NN $X \to b\bar{b}$ tagger became available in 2020. And the first ever ATLAS full calibration for this $b\bar{b}$ tagger followed in 2021, and it was object of a public ATLAS note [30]. The fruit of a collaboration between University of Bonn, LPNHE/CNRS and Humboldt Universität zu Berlin. The $Z(\to b\bar{b})\gamma$ calibration effort followed closely the methodology and selection established in previous works exploring this topology (see previous paragraph). LPNHE led the $Z(\to b\bar{b})$+jets calibration effort, with major contributions from Yajun He, a

PhD student at LPNHE who I co-supervise between 2019 and 2022.

4.3.1 Boosted $X \to b\bar{b}$ tagger signal calibration

The calibration of the tagger is performed according to a process-based approach, as opposed to the flavour-based approach typically pursued in the calibration of small-R jets discussed in Chapter 1. The signal efficiency of the tagger is calibrated using $Z(\to b\bar{b})\gamma$ events for $200 < p_{\mathrm{T}} < 450$ GeV due to statistics limitations and using $Z(\to b\bar{b})$ + jets for $p_{\mathrm{T}} > 450$ GeV. The calibration and the modelling study was carried out for the 60% signal efficiency working point and for four large-R jet p_{T} bins between 200 and 1000 GeV. Details about the MC simulation samples used and the handling of systematic uncertainties can be found in Ref. [30].

The data-to-simulation scale factors applied to MC to match the tagger signal efficiency to the one measured in data can be calculated as:

$$\mathrm{SF} = \frac{\epsilon^{\mathrm{data}}}{\epsilon^{\mathrm{MC}}} = \frac{\frac{N^{\mathrm{data}}_{\mathrm{passed}}}{N^{\mathrm{data}}_{\mathrm{total}}}}{\frac{N^{\mathrm{MC}}_{\mathrm{passed}}}{N^{\mathrm{MC}}_{\mathrm{total}}}} = \frac{\frac{N^{\mathrm{data}}_{\mathrm{passed}}}{N^{\mathrm{MC}}_{\mathrm{passed}}}}{\frac{N^{\mathrm{data}}_{\mathrm{total}}}{N^{\mathrm{MC}}_{\mathrm{total}}}} = \frac{\mu_{post-tag}}{\mu_{pre-tag}}. \tag{4.1}$$

The suffix passed refers to the number of data or MC events which pass the $X \to b\bar{b}$ tagger requirement and total refers to the total number of signal events in data or MC. The values $\mu_{pre-tag}$ and $\mu_{post-tag}$ describe the number of signal events in data divided by the number of signal events in MC simulation (referred to as the signal strength) before and after tagging, respectively.

To measure $\mu_{post-tag}$, events with at least one large-R jet identified as a Z-boson candidate are selected and the mass of the leading p_{T} large-R jet is used to determine the signal strength in a likelihood fit. As already mentioned, two processes with $Z \to b\bar{b}$ were used to measure both signal strength parameters:

$Z+\gamma$: The trigger single photon transverse energy threshold was $E_{\mathrm{T}} > 120$ GeV in 2015 and increased to 140 GeV in 2016-2018 [80].

The leading-p_{T} large-R jet fulfilling the following requirements is identified as the Z-boson candidate: $p_{\mathrm{T}} > 200$ GeV, a jet mass above 50 GeV, $2m_{Z\to b\bar{b}}/p_{\mathrm{T}} < 1.0$ to ensure that Z-boson candidate decay products are contained inside the $R = 1.0$ cone, contains at least two ghost-associated VR track-jets with $p_{\mathrm{T}} > 7$ GeV and pass the $X \to b\bar{b}$ tagger requirement at the 60% efficiency WP. The $Z-$boson candidate is further required to be well separated from the photon ($\Delta\phi(\text{jet},\gamma) > \pi/2$). The main background in the $Z(\to b\bar{b})\gamma$ measurement is γ+jets. Contributions from Z+jets, where a jet is mistakenly reconstructed as a photon, in addition to the $W\gamma$, and $t\bar{t}\gamma$ processes, are small.

Z+jets: The analysis uses unprescaled single large-R jet triggers [118, 121] which evolved over the Run-2 data-taking periods to adjust to machine conditions by increasing the jet p_{T} threshold and introducing a jet mass requirement. Events are required to have at least two large-R jets with $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.0$ and no electrons or muons with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Additional requirements are applied to the large-R jets to suppress the mis-modelling of multijets and to reject backgrounds, these cuts are inspired on the hadronic VV search analysis presented in Sections 2 and 3:

- $\frac{p_{\mathrm{T},1}-p_{\mathrm{T},2}}{p_{\mathrm{T},1}+p_{\mathrm{T},2}} < 0.15$, where $p_{\mathrm{T},1}$ ($p_{\mathrm{T},2}$) is the transverse momentum of the (sub-)leading large-R jet,
- $|\Delta y_{1,2}| < 1.2$, the rapidity difference of the leading and sub-leading large-R jets.

The leading-p_{T} large-R jet is required to fulfill the requirement $p_{\mathrm{T}} > 450$ GeV (due to the trigger selection), a jet mass above 50 GeV, to contain at least two ghost-associated VR track-jets with $p_{\mathrm{T}} > 7$ GeVand to pass the $X \rightarrow b\bar{b}$ tagger requirement at the 60% efficiency WP to be identified as the Z-boson candidate. The dominant source of background in the $Z(\rightarrow bb\bar{b})$+jets measurement is multijet events, with small contributions of $W(\rightarrow q\bar{q}')$ + jets and $t\bar{t}$ events.

To measure $\mu_{post-tag}$, the signal region is defined by a mass window requirement of $50 < m_{Z\rightarrow b\bar{b}} < 150$ GeV of the tagged Z-boson candidate.

The use of $Z \rightarrow b\bar{b}$ events to measure $\mu_{pre-tag}$ is extremely difficult due to the overwhelming background before tagging. Therefore $\mu_{pre-tag}$ is obtained using the $(Z \rightarrow \ell^+\ell^-)$+jets and $(Z \rightarrow \ell^+\ell^-)$+$\gamma$ processes, which exhibits a substantially smaller relative background contribution. The number of $Z \rightarrow b\bar{b}$ events in data before tagging can be estimated at the first order as:

$$N^{\mathrm{data}}_{Z\rightarrow b\bar{b}} = N^{\mathrm{MC}}_{Z\rightarrow b\bar{b}} \cdot \frac{N^{\mathrm{data}}_{\ell\ell} - N^{\mathrm{MC}}_{\mathrm{bkg},\ell\ell}}{N^{\mathrm{MC}}_{Z\rightarrow \ell^+\ell^-}}, \tag{4.2}$$

where $N^{\mathrm{MC}}_{Z\rightarrow b\bar{b}}$ is the number of $Z \rightarrow b\bar{b}$ signal events in MC before tagging, $N^{\mathrm{data}}_{\ell\ell}$ is the total number of selected events in data, $N^{\mathrm{MC}}_{\mathrm{bkg},\ell\ell}$ is the number of background events predicted by MC simulations and $N^{\mathrm{MC}}_{Z\rightarrow \ell^+\ell^-}$ is the number of $Z \rightarrow \ell^+\ell^-$ signal events predicted by MC simulation. Thus, $\mu_{pre-tag}$ can be calculated as:

$$\mu_{pre-tag} = \frac{N^{\mathrm{data}}_{\ell\ell} - N^{\mathrm{MC}}_{\mathrm{bkg},\ell\ell}}{N^{\mathrm{MC}}_{Z\rightarrow \ell^+\ell^-}}, \tag{4.3}$$

and depends only on the yields obtained in the $Z \rightarrow \ell^+\ell^-$ channel. The $Z \rightarrow \ell^+\ell^-$ selection closely follows the kinematic selection used for $\mu_{post-tag}$ measurements in order to ensure compatibility in the definition of the Z-boson candidate.

The multijet and γ+jets backgrounds are determined using data-driven techniques, using classes of exponential and polynomial functions of 2nd to 5th order. Other backgrounds and signal are modelled using MC simulations. Techniques differ for the $\mu_{pre-tag}$ and $\mu_{post-tag}$ measurements. For the $\mu_{post-tag}$ measurement a likelihood fit to the $Z \rightarrow b\bar{b}$ mass distribution is performed, while for the $\mu_{pre-tag}$ measurement only signal and background yields need to be extracted. Figure 4.8 shows the shows the $Z \rightarrow b\bar{b}$ candidate invariant mass distribution after the fit and applying the $X \rightarrow b\bar{b}$ tagger requirement at 60% working point. The pulls, defined as data minus fitted model prediction divided by the data statistical uncertainty, are mostly within three standard deviations.

The SFs are measured as a function of the Z-boson candidate p_{T}. Four bins are defined based on the number of events observed in data: $200 < p_{\mathrm{T}} < 450$ GeV, $450 < p_{\mathrm{T}} < 500$ GeV, $500 < p_{\mathrm{T}} < 600$ GeV and $600 < p_{\mathrm{T}} < 1000$ GeV. The SF in the first p_{T} bin is measured using the $Z(\rightarrow b\bar{b})\gamma$ events and the SFs in the last three p_{T} bins are measured using the $Z \rightarrow b\bar{b}$ +jets events. The scale factors are illustrated in Figure 4.9. The main systematic uncertainties for the $Z(\rightarrow b\bar{b})\gamma$ selection are the Z-boson modelling, the spurious signal and the jet energy scale. The statistical uncertainty represents around 65% of the total error. In the case of the $Z \rightarrow b\bar{b}$ jets selection, the main systematic uncertainties are Z+jets modelling, the uncertainty on the fit model and jet energy and mass resolution/scale; while the statistical uncertainty represents around 30-35% of the total error, depending on the Z-boson candidate p_{T}.

These calibrations are now available to be used by ATLAS Run-2 analyses using boosted $H \rightarrow b\bar{b}$ topologies and more generally boosted $b\bar{b}$ topologies since the $X \rightarrow b\bar{b}$ tagger's large-R jet mass correlation is very small [30].

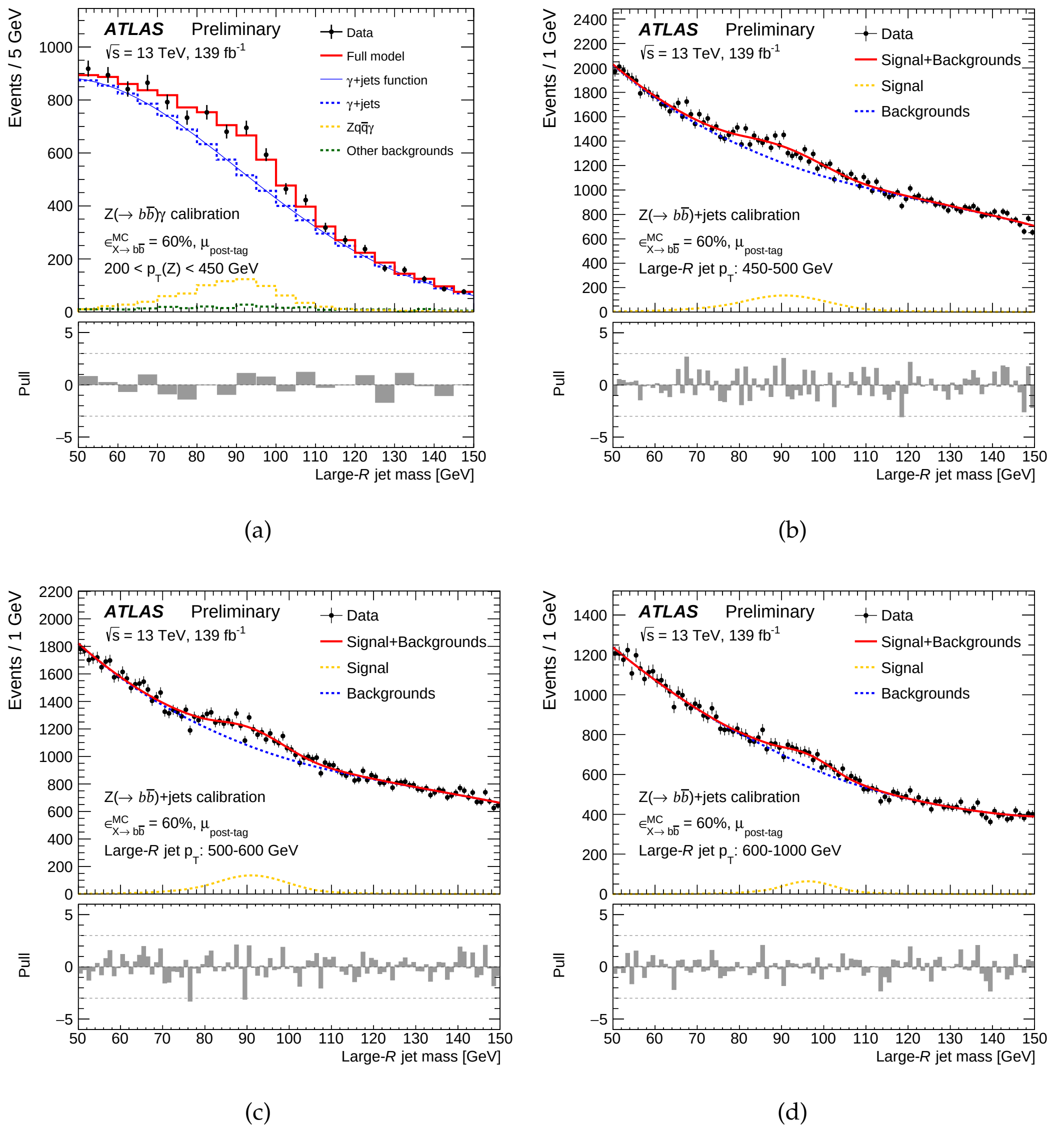


FIGURE 4.8: The $Z \to b\bar{b}$ candidate invariant mass distribution after the fit and applying the: (a) $Z(\to b\bar{b})\gamma$ selection for $200 < p_{\mathrm{T}} < 450$ GeV and (b,c,d) $Z \to b\bar{b}$ +jets selection and the $X \to b\bar{b}$ tagger requirement at 60% working point for events with the large-R jet (b) $450 < p_{\mathrm{T}} < 500$ GeV (c); $500 < p_{\mathrm{T}} < 600$ GeV and (d) $600 < p_{\mathrm{T}} < 1000$ GeV range [30].

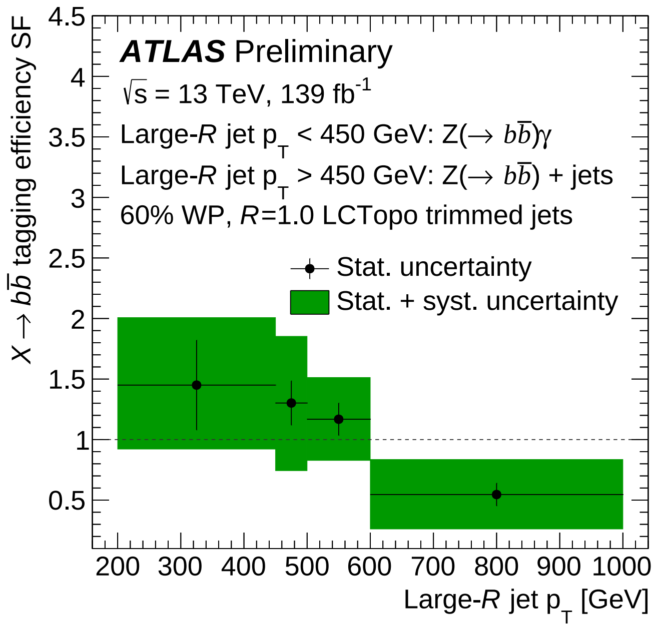


FIGURE 4.9: Signal efficiency scale factors for $X \to b\bar{b}$ tagger at 60% efficiency working point [30].

4.4 Summary and SM Higgs global context

This chapter briefly described the first measurement of the Higgs boson production in association with a W or Z boson in the $b\bar{b}$ decay channel at high transverse momentum with the ATLAS experiment at $\sqrt{s} = 13$ TeV. Preparations for the legacy Run-2 boosted $H \to b\bar{b}$ analyses were also discussed. The observation of the $H \to b\bar{b}$ channel and the start of its precision measurements era was one of the major results in the Higgs boson saga in Run-2. But there are many more. Since its discovery in 2012 more than 30 times as many Higgs bosons have been recorded by the ATLAS and CMS experiments, each, allowing to improve the precision of the measurements and have access to many decay channels and production modes. Globally this is the current status:

- All main production and decay modes have been observed
- First hints of rare Higgs decays have been seen ($H \to \mu\mu$)
- Total and differential cross sections have been measured and used to extract information on the Higgs couplings to fermions and vector bosons and put constraints in EFT models (and therefore on indirect searches for new physics)
- Kinematic dependence of production cross sections has been studied in particular for the discovery channels: $H \to \gamma\gamma, WW, ZZ$
- Progress on setting the baseline for di-Higgs studies

For the 10 years anniversary ATLAS performed an ambitious combination of the latest SM Higgs analyses including the following channels targeting several production modes: $\gamma\gamma, ZZ, WW, Z\gamma, b\bar{b}, \tau\tau, \mu\mu, c\bar{c}$ and invisible [9]. Almost all measurements included used the full Run-2 dataset collected by the experiment. The global signal strength measurement has an SM compatibility of 39%: $\mu = 1.05 \pm 0.06 = 1.05 \pm 0.03(stats.) \pm 0.03(exp.) \pm$

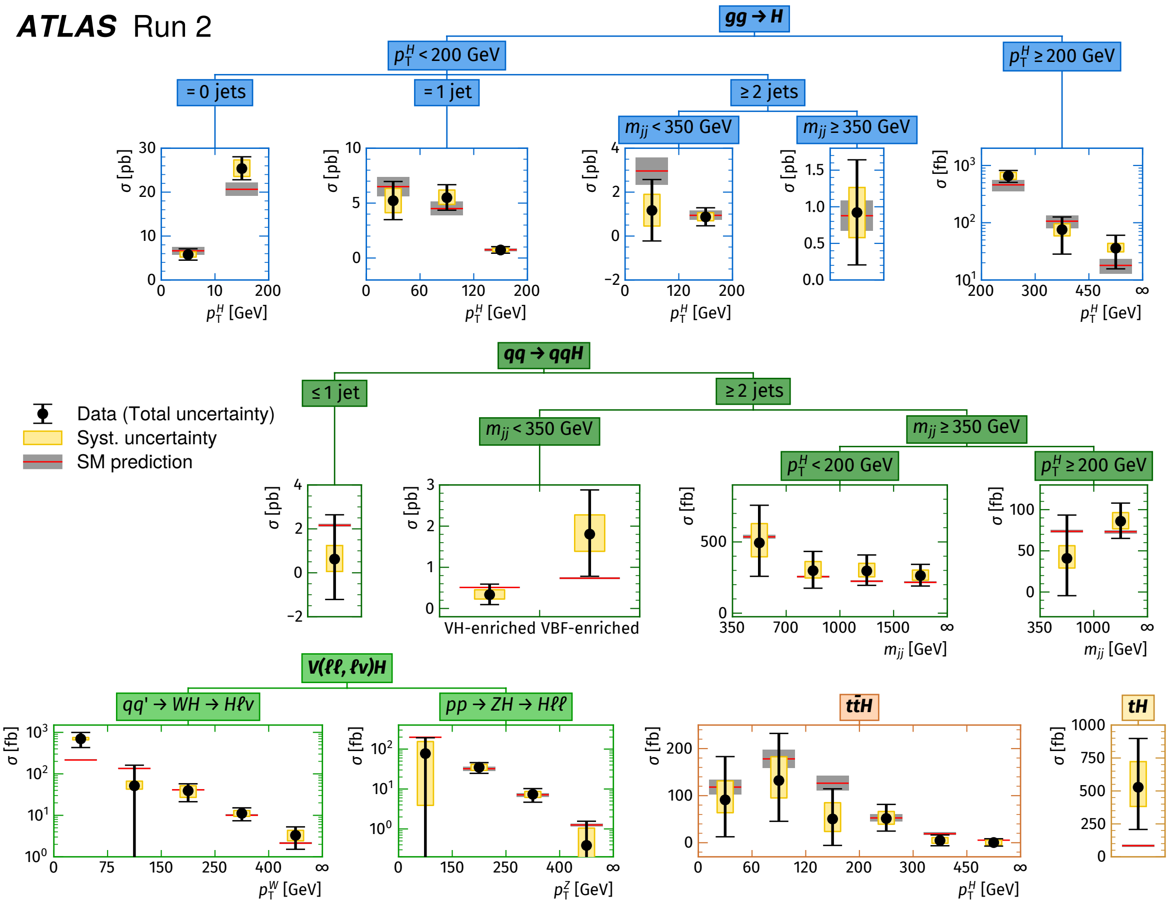
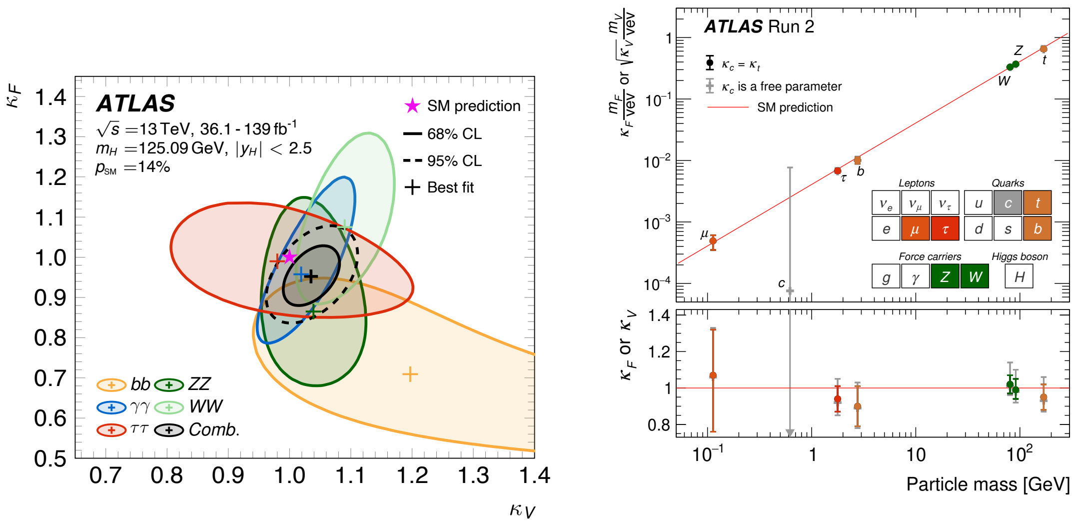


FIGURE 4.10: Measured STXS cross-sections in different kinematic regions for the latest ATLAS global Higgs analyses combination performed [9].

$0.04(sig.th.) \pm 0.02(bkg.th.)$. It is worth mentioning that experimental and theory uncertainties are half of the ones obtained in Run-1 combinations thanks to three main raisons: more statistics, better upgraded detectors, extensive use of machine learning and improved detector objects reconstruction and calibration despite de high Run-2 pile-up conditions. STXS cross-sections were measured in 36 kinematic regions and the results are shown in Figure 4.10.

The results for the Higgs couplings are shown in Figure 4.11. The κ framework [199] was used in this case, where coupling modifiers κ are used to quantify the Higgs couplings deviations from the SM predictions. In general good compatibility with the SM was found. Previous ATLAS and CMS combinations also interpreted the results in terms of EFT-based frameworks. The global move to constrain EFTs rather than signal strengths for couplings represents a new, more comprehensive phase in SM tests at the LHC and the full set of results for the previous combinations can be found in Refs. [22, 165].

There has been great progress in understanding the Higgs boson mechanism since its discovery. A decade of measurements has strengthened the hypothesis that its properties are in agreement with the BEH predictions but the full story is not known yet (more in next Chapter) and exciting times ahead with the start of Run-3, the HL-LHC phase and possible future colliders!



(a) (b)

FIGURE 4.11: (a) Log-likelihood contours for the plane of modifier for vector bosons κ_V and for fermions κ_F for the individual decay modes (colours) and their combination (black). They are obtained assuming positive coupling strengths to fermions and bosons, and no contributions from invisible or undetected non-SM Higgs boson decays. The best-fit value for each measurement is indicated by a cross while the SM hypothesis is indicated by a star. (b) Higgs boson coupling strength modifiers and their uncertainties as a function of their masses. Loop-induced processes are assumed to have the SM structure, and Higgs boson decays to non-SM particles are not allowed [9].

Chapter 5

Conclusion and outlook

The particle content of the Standard Model was completed in 2012 with the discovery of the Higgs boson. A decade of measurements has strengthened the hypothesis that the Higgs properties are in agreement with the electroweak symmetry breaking mechanism in the SM, the BEH predictions. However, an important number of fundamental physics questions remain unanswered and due to its exceptional role in the SM, the study of the EWSM mechanism and the Higgs boson itself might be the path to new physics. This document explored several routes to discovering new physics using diboson topologies (including vector bosons W, Z and Higgs bosons), so tightly related with the EWSM mechanism. First, we described the direct search for new resonances decaying to VV and VH, with $V = W, Z$ and considered the use of boosted boson tagging methods to extend the sensitivity of these searches to higher resonance masses. Second, we delved into indirect searches of new physics, looking for data deviations from the simulation prediction in the vector boson associated production of the Higgs boson, $VH, H \to b\bar{b}$. We reviewed how measurements of the VH production can be optimised in the high transverse momentum regime following the use of boosted boson tagging methods and set constraints on new physics using the EFT frameworks.

At the end of each chapter we discussed the state of art for each one of these studies, how they situated in the global physics program at the LHC and provided hints of near future possible improvements. **In this chapter we will briefly explore the longer-term expectations for future LHC runs, as well as interesting new studies to lead the way towards a better understanding of the electroweak symmetry breaking mechanism using dibosons**. The troy horses that will be needed are: **more data, improved analysis strategies and objects reconstruction and calibration techniques and detector upgrades** adapted to the harsh LHC's collision environment.

5.1 More data

The LHC plan includes operations for another 20 years (see Figure 5.1). The analyses presented in this document used data from Run-1 and Run-2. During Run-3, started during the summer of 2022, the integrated luminosity in pp collisions should reach between 200 and 300/fb before a long shutdown that will prepare the accelerator for a high-luminosity running phase (HL-LHC). The HL-LHC should deliver 3000 fb^{-1} in ten years with peak luminosities of 5-7.5$\times$10^{34} cm^{-2} s^{-1} and up to 200 collisions per bunch crossing.

The dataset expected to be collected in Run-3 and HL-LHC, will increase the sensitivity of the analyses presented here. Extrapolations dating back to 2019 (and based on early analysis techniques and object reconstruction performance) reported expected upper limits on the signal cross section times BR for HVT W' and Z' to be 4.3 TeV with 300/fb and 4.9 TeV with 3/ab of pp collisions [148] for the $VV \to \ell\nu qq$ channel. For the Bulk graviton the expected limits were estimated as 2.8 and 3.3 TeV at 300/fb and 3/ab [148]. The values at 3/ab showed an expected increase to the sensitivity of the search to the benchmark signals

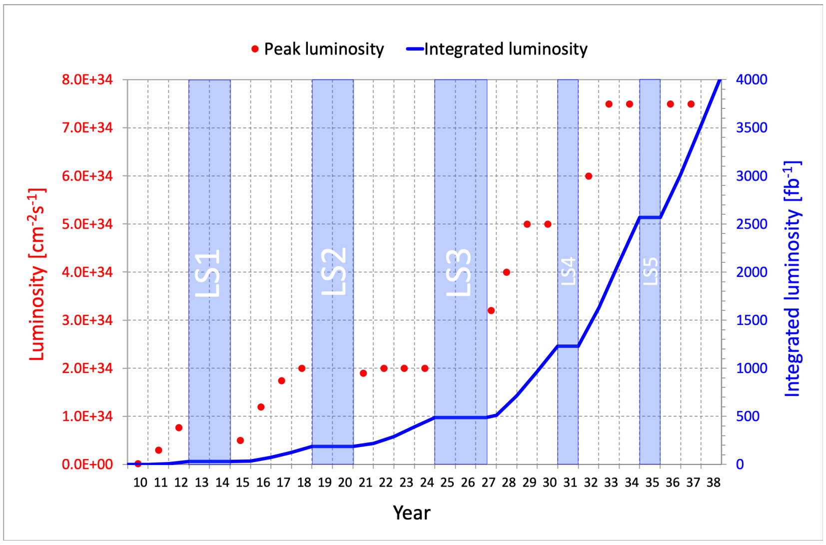


FIGURE 5.1: Long-term LHC schedule, showing the expected evolution of the peak and integrated luminosities in pp collisions, as well as the long shut-down (LS) periods. From Ref. [201].

by 1 TeV with respect to the limits in the semileptonic channel in the early Run-2 analyses presented in this document. Searches in the $\ell\ell qq$ and fully hadronic decay channels are expected to have similar sensitivities at high masses, following the observation in the ATLAS searches with Run-2 data.

Regarding the VH SM analysis, with 3000/ab collected during HL-LHC a dataset of around 200 million Higgs bosons will be produced, from which more than half will be decaying via $H \to b\bar{b}$. This large dataset should allow a precise study of the kinematics of Higgs boson production and decays. Extrapolation performed from $VH, H \to b\bar{b}$ results obtained with the 2015-2017 80/fb datasets show an improvement of the signal strength measurement precision by a factor 2.5-3 in the uncertainty of the measurement [145]. Moreover, more precise measurements of kinematic observables such as the higher transverse momentum can probe possible deviations induced by new phenomena [145].

5.2 New physics analysis strategies

A decade of measurements of Higgs boson properties confirms the existence of this scalar field, but its doublet structure and the shape of the associated potential remain largely unknown. The BEH mechanism implies the existence of a Higgs boson self-interaction, HHH, and its interactions with vector bosons, VVH and $VVHH$, whose strengths are tightly interconnected in the model. The study of the di-Higgs boson gluon-fusion and VBF production will be fundamental to understand the nature of the Higgs boson and of EWSB and discover or tightly constrain BSM models. This topic is very interesting to me for two reasons. First, the analysis of the full Run-2 and Run-3 data sets and the use of boosted Higgs taggers will increase the experimental sensitivity on di-Higgs production cross section. Second, a global combination including diboson, single Higgs and di-Higgs inputs will improve the current constraints on EFT interpretations. As the underlying dynamics of new physics is not really known and the impact of new interactions with SM particles can be subtle, it is important to combine as many measurements as possible across the full spectrum of the LHC research programme.

In the meantime and while waiting for the Run-3 and HL-LHC, another approach to improve the current results is to diversify the Run-2 analysis to cover non-explored channels and new corners of the phase-space. In spring 2021, I got involved in the development of an analysis measuring the fully hadronic VH, $H \to b\bar{b}$ process using the dataset accumulated at Run-2 with an a PhD student at UPC/LPNHE I am co-directing, Yajun He. The analysis targets the high transverse momentum regime and uses large-R jets, jet-substructure information and the NN $X \to bb$ tagger calibrated as part of Yajun's qualification task, described in Section 4.3. The analysis is made in collaboration with teams from SLAC, BNL, University of Massachusetts and Santa Cruz Institute of Physics. Its large cross-section bring the potential to reach higher transverse Higgs momentum but it comes with the challenge of a large multijet background. It will then be interesting to combined this analysis with the semileptonic VH, $H \to b\bar{b}$ topologies in order to understand their complementarity in the parameter space. This will allow us to better prepare the analyses and reconstruction algorithms for Run-3 and beyond.

5.3 Improved object reconstruction

Increasing the sensitivity of these analyses can also be achieved through objects performance improvements, increasing the identification efficiency of certain objects for a given level of background rejection, for example, or through the reduction of certain systematic uncertainties (some reduction scenarios are already taken into account in the prospective studies mentioned above). Some examples of performance improvements were already mentioned in previous chapters, i.e. the reduction, by a factor of about 2, of the uncertainty on the energy scale of large-R jets, thanks to the use of in situ calibration techniques in Section 1.4.4 or the new UFO W/Z taggers defined using machine learning techniques [87] mentioned in Section 2.4.

As one of the current conveners of the Jet/Etmiss group in ATLAS (started in October 2021) my focus in the next couple of years will be on scrutinizing and providing support to a large number of performance analyses and new developments, such that good quality results can be obtained in a timely manner. Two of the main goals of the group work in 2022 and 2023 will be to finalize the jet inputs, calibration and tagging recommendations for Run-2 legacy physics analyses and to update those recommendations for the first physics analyses using Run-3 data. Beyond my current role within the jetEtmiss group, I continue my involvement in boosted hadronic boson taggers development, their calibration and derivation of corresponding uncertainties. Current multivariate boosted boson taggers using ML are based on high-level jet substructure and b-tagging information. Recent studies have shown an important performance improvement of such taggers when using the kinematic properties of the jet's constituents instead, like ParticleNet [23, 208] and Lund-jet plane [172, 173]. This is possible since the models can access a superset of the information not contained in any set of high-level quantities. For example, the Lund-jet plane is the two-dimensional space of the momentum fraction and opening angle for the gluon emissions relative to the primary quark or gluon core inside a jet, so it directly accesses the jet components information instead of high-level information as jet substructure variables. Such tagger can connect directly with individual regions of the phase space, facilitating also the evaluation of associated systematic uncertainties.

5.4 Upgrade our detectors

Another important ingredient to increase the sensitivity of the analyses that constitute the LHC physics program and to even extend such program is the instrumentation R&D. The high luminosity foreseen for the HL-LHC implies a much harsher environment than in Run-2 for the ATLAS sub-detectors. Indeed high luminosity means higher event rate, more pile-up events and higher radiation doses and fluences. To cope with the severe data taking conditions expected at the HL-LHC several upgrades to the ATLAS detector are planned. The LPNHE group is committed to one of these upgrades: a complete new silicon only tracker called the Inner Tracker (ITk). ITk will have to not only guarantee the same (or better) level of performance of the current ATLAS Inner Detector but also expand the tracking coverage beyond the current $|\eta| = 2.5$ to $|\eta| = 4.0$ (see Figure 5.2). A good tracker will be essential for b-tagging and jet reconstruction in the HL-LHC era. Since early 2018 I got involved in the ITk group activities, actively contributing to the silicon pixel module assembly and characterisation effort. My activities were mainly related with gluing, metrology and support for measurement campaigns on test-beam at CERN and DESY. LPNHE in collaboration with CEA-Saclay and IJClab have engaged to participate in the pixel module production phase that will start in 2023-2024. Beyond that period I hope to have the opportunity to exploit this detector and its performance in terms of tracking, vertexing and flavour tagging for precision measurements and discoveries.

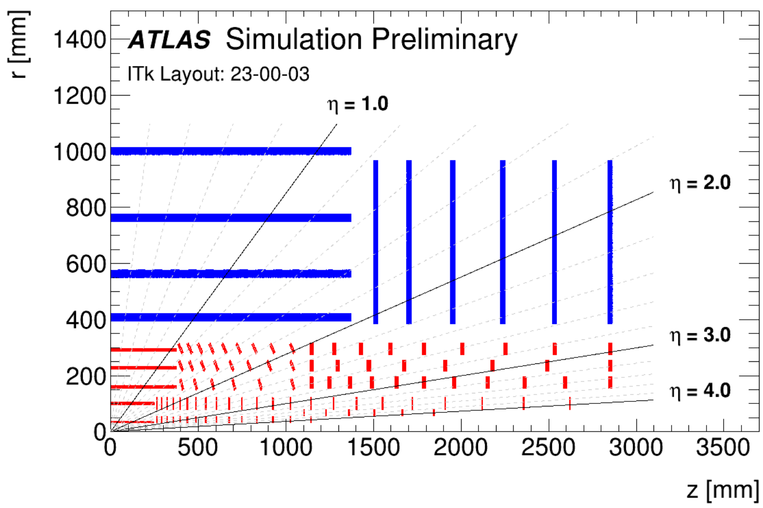


FIGURE 5.2: A schematic representation of the new internal detector proposed for the HL-LHC as a function of the distance to the interaction point along the beam axis (distance z) and the radius in transverse plane (distance R). From Ref. [37].

5.5 Before closing

The analyses, performance and detector activities presented in this document and the ongoing ones mentioned in this last chapter are vital steps in a long journey towards measuring the properties of the SM particles like the Higgs boson exploring the possibilities of new physics. At the time of writing, all the data accumulated at Run-2, i.e. 139 fb-1 of data at 13 TeV, is being analysed and Run-3 just started a couple of months ago. Many results should therefore be made public in the coming months. Run-3, which will multiply the data by a factor of about two, and the HL-LHC, which will increase the data set tenfold, will eventually increase the sensitivity of existing analyses.

As physicists we enhance our algorithms, improve our understanding of background processes, upgrade our detectors and collect more data, venturing ever further into uncharted territory where new physics may await. We do it, while creating new and strengthening already existing scientific collaborations. We do it while training the new generations of scientists to the best of our abilities. I strongly believe that scientific and educational capacity building, for academia and beyond it, as well as the connection with the wider society are fundamental to us, scientists, to achieve high quality, international and responsible science based on principles of equity, diversity and inclusion.

We started the introduction to this document with the following statement "we hope new information from experiments at the LHC will help us to find more of these missing pieces" and indeed that's certainly my hope, but what if that's not the case? Exploratory studies are necessary to ensure a solid scientific case is made for future experiments. Hopefuly the LHC and HL-LHC will help us to address some of the current open questions in the field and to define the strategy for the next facility to explore the energy frontier in particle physics.

Appendix A

Teaching and student supervision

A.1 Teaching

- 02/2021-06/2021 and 02/2022-06/2022: «Scientific Instrumentation» course for LA-CoNGA physics. Common course for 8 universities in Colombia, Ecuador, Peru and Venezuela (2 semesters). More information about the course can be found at «Scientific Instrumentation» syllabus

- 2019-2020: « Physics I » for 1st year engineering students. IPSA- Ecole d'ingénieurs aéronautique et spatiale, Paris, France (3 semesters)

- 09/2014-09/2017: Virtual course of introduction to particle physics for first year master students in physics. It is the first course of this type taught in Venezuelan, Colombian and Peruvian universities. The notions of Standard Model, principles of detectors and data analysis in particle physics are covered. This course was held once a year between 2014 and 2018 with 10-15 students per year. I participated in in the general organization of the course, and more precisely I was in charge of the detectors section. More information can be found at (in Spanish): http://www.cevale2ve.org (4 semesters)

- 01/2013-11/2014: Postdoctoral teaching assistant (TD), «Electrodynamics» for 1st year bachelor physics students. Université de Genève, Geneva, Switzerland (3 semesters)

- 02/2005-02/2008: Undergraduate teaching assistant (TD/TP), « Mathematical methods for physicists» (linear and multilinear algebra, complex analysis) for 3rd years physics bachelor and Kinematic, Mechanics and Optics laboratory for 1st and 2nd year physics and engineering bachelor students. Universidad de Los Andes, Mérida, Venezuela.

A.2 Student supervision

A.2.1 PhD level

- Yajun He. Co-directing with Giovanni Calderini. Started her thesis in October 2019: « Identification of boosted higgs bosons for searches of new physics and improvement of the ATLAS tracker for high luminosity phase of LHC ». Thesis defense expected in Autumn 2022.

During my time (and until mid-2018) as a postdoctoral researcher at the University of Chicago within David Miller's group, I contributed to the supervision of the following students:

- Joakim Olsson, « Searching for Supersymmetry in Fully Hadronic Final States with the ATLAS Experiment ». Worked together on a search for electroweak production of a chargino-neutralino pair, decaying into SM-quarks via a W boson and a SM-like 125 GeV Higgs boson, performed using an integrated luminosity of 36.1/ fb at 13 TeV, Phys.Rev. D100 (2019) no.1, 012006. Defended his thesis in September 2018. Currently postdoc at UC Irvine, US.

- Giordon Stark, « The search for supersymmetry in hadronic final states using boosted object reconstruction ». Worked together on global Feature EXtractor (gFEX) performance studies between 06/2015 and 12/2017. gFEX is a project of the L1Calo phase-I upgrade of the ATLAS detector. Defended his thesis in April 2018. Currently postdoc at UC Santa Cruz, US.

- Miles Wu, « Measurement of Collinear W Boson Emission from High Transverse Momenta Jets Using the ATLAS Detector ». Worked together on the measurement of the jet mass in high transverse momentum $Z(b\bar{b})$ production at 13 TeV using the ATLAS detector, Phys. Lett. B 812 (2020) 135991 between 06/2015 and 12/2016. Defended his thesis in March 2017. Currently working on financial services.

During my time as postdoctoral researcher at the University of Geneva, within Giuseppe Giacobucci's group, I contributed to the thesis supervision of Chris Delitzsch. Supervision was then continued by a colleague, Steven Schramm, and we continued collaborating in several projects afterwards.

- Chris Delitzsch, « Search for Diboson Resonances in the Fully Hadronic Final State Using Jet Substructure Techniques in 8 and 13 TeV Proton-Proton Collisions with the ATLAS Detector ». Worked together on several topics during her thesis, including but not limited to jet calibration and tagging and the search for hadronically decaying VV resonances at 8 and 13 TeV between 01/2013 and 12/2018. Defended her thesis in December 2016. Currently working as postdoctoral researcher at Technische Universitaet Dortmund, Germany.

A.2.2 Master level

- Rafael Vinasco Soler, Universidad Nacional de Colombia, in the context of the LA-CoNGA physics program: « Development of a boosted boson tagger using Lund Jet plane information », March 2021-Present. Supervisors : Reina Camacho Toro, Mykola Khandoga (LPNHE/CNRS), Carlos Sandoval (Universidad Nacional de Colombia, Colombia)

- Yajun He, Université Paris Diderot, in the context of the NPAC (Nuclei, Particles, Astroparticles and Cosmology) M2 master: « Identification des bosons de Higgs boostés pour la nouvelle physique », May-June 2019. Supervisor: Reina Camacho Toro

- Raphaël Bajou, Université Paris Diderot, M1: « Performance of pixel sensors before and after irradiation using testbeam data taken at Fermilab, Chicago», June-July 2018. Supervisors: Reina Camacho Toro and Marco Bomben (LPNHE/CNRS). Raphaël fait actuellement son doctorat dans l'APC Paris, dans l'experience DUNE.

A.2.3 Bachelor level

- Mariana Vivas, Universidad Central de Venezuela: « Comparative analysis of ML techniques for new physics searches in dijet events using open data from the LHC

Olympics: a reproducible and open study », September 2020-October 2022. Supervisors: Reina Camacho Toro, Camila Rangel (The Turing Institute, UK), José Antonio Lopez (Universidad Central de Venezuela, Venezuela)

- Laura Caviedes Betancourt and Juan David Villamil Santiago, Universidad Nacional de Colombia, in the context of the LA-CoNGA physics program: « First look at the CI/CD workflow for the event selection of Z+jets calibration ». March-July 2021. Supervisors: Reina Camacho Toro, Yajun He, Carlos Sandoval (UNAL).

- Muriel Vilage, Université Paris Diderot, L3: « Mesures de grandeurs électriques à différentes températures de diode en silicium irradiées ayant pour but d'étudier la dépendance de la température au courant de fuite ». June-July 2019. Supervisors : Marco Bomben (LPNHE/CNRS) and Reina Camacho Toro

- Sarah Kroeker, in the context of the summer student program from Fresno State University, US (based at CERN): « Development of jet reconstruction algorithms for gFEX ». July-September 2016. Supervisors: Reina Camacho Toro and Giordon Stark.

- Claudia Clarke, in the context of the summer student program from Université de Genève, Switzerland (based at CERN): « Study of the Higgs boson decay into b quarks in the boosted regime ». July-September 2014. Supervisor: Reina Camacho Toro

- Loic Chagot, l'Université Blaise Pascal, Clermont Ferrand in the context of the JANUS program: « Étude de la désintégration du quark top ». July-August 2011. Supervisor: Reina Camacho Toro.

A.3 Teaching and students supervision training followed

- Echanges de pratiques entre encadrants de doctorants. Association Bernard Gregory (ABG). France. April 2020 (2 hours).

- Management d'un projet doctoral. Institut de formation doctorale, UPMC-Sorbonne Universit'es. 2019 (2 days).

- Conseil collectif d'encadrants. Sorbonne Université. 2019 (4 half days).

Appendix B

Other responsabilities

B.1 Within the ATLAS Collaboration

- 10/2021-Present, convener of the jetEtmiss group. 2 years position. Two of the main goals of the jetEtmiss group work in 2022 and 2023 will be to finalize the jet inputs, calibration and tagging as well as MET recommendations for Run-2 legacy physics analyses and to update those recommendations for the first physics analyses using Run-3 data. Currently the group is composed of around 100 people.

- 10/2019-09/2021, convener of the Flavour Tagging Calibration subgroup. I was responsible for the coordination and supervision of all results of this subgroup related to the calibration of the b-tagging algorithms used by the collaboration. At the time of the convenership, the subgroup included about 10 analyses and about 30 people. During my convenership 2 papers, 4 conference notes and 2 rounds of recommendations for ATLAS analysers were produced using Run-2 data.

- 03/2019-02/2021, Member of the Early Career Scientist Board (ECSB) of ATLAS. The ECSB is a special advisory group dedicated to helping the ATLAS collaboration create an environment where the full scientific potential of young scientists can be realized.

- 04/2016-04/2017, convener of the Dibosons, Multileptons and Extra Dimensions (DBL) subgroup belonging to the Exotics group. I was responsible for the coordination, supervision and pre-approval of all results of this subgroup. The main goal of this subgroup was to explore the search for diboson and leptonic resonances in different topologies/decays. During my mandate, I guided and supervised 10 different analyses, 7 of them were the subject of public notes presented in conferences. The group was composed of around 50 members.

- 01/2015-07/2016, coordinator for the $X \to b\bar{b}$ performance taskforce, a team at the time under the supervision of the ATLAS Jet Substructure and Tagging Performance and Flavour Tagging Groups. A team of around 15 members working on the optimisation and calibration of tagger for $b\bar{b}$ topologies, targeting in particular Higgs and Z boson decays.

- Publication editor, member of editorial boards on the jetEtmiss , Exotics and Higgs groups.

- Member of search committees for Publications Committee members (2018), ECSB members (2020) and the conveners of the physics groups (2021)

B.2 Organization of workshops, conferences and other events

- ATLAS Hadronic Calibration Workshop (HCW). Scientific Committee. Convener of the ATLAS jetEtmiss group. August 30-September 2, 2022. Cambridge, UK.
- Member of the organizing committee of the 1st LA-CoNGA physics Network School. Bucaramanga, Colombia. December 2021
- Member of the local organizing committee for the Higgs Hunting 2021 and 2022 conference. Online. September, 2021
- Member of the scientific organizing committee of the « Physique ATLAS France » workshop 2021. Online. June-July 2021
- Organizer of the panel "Open science, shared infrastructure and collaborative networks: the challenge of the new higher education" at the Open Forum of Sciences of Latin America and the Caribbean – CILAC 2021. Online. April 2021.
- Organizer of the session "Accelerators, Detectors, Instrumentation and Computing in Latin-America", II Latin American Strategy Forum for Research Infrastructure (LASF4RI) : an Open Symposium for HECAP, Online, juillet 2020
- Chair of the Electroweak session, ATLAS Standard Model Workshop 2019, Belgrade, Serbie. Septembre 2019
- Member of the local organising committee for the LPNHE Biennale meeting, Montpellier, France. May 2018.
- Member of the local organizing committee for the International Workshop on Boosted Object Phenomenology, Reconstruction and Searches in HEP (BOOST) in Paris July 2018 and Chicago in August 2015, and member of its international scientific committee since 2019.
- Co-organiser of the LPNHE laboratory weekly meetings (known as «réunion du Vendredi») (2018-2020). Also in collaboration with doctoral students and postdocs of the laboratory I have set up the Journal Club of the laboratory in 2018.
- Organiser of the session « Standard Model » Journées de Rencontre des Jeunes Chercheurs 2018 (Lège-Cap-Ferret, France, octobre 2018) and 2019 (Logonna-Daoulas, France, novembre 2019).
- Chair of the session « Supersymmetry: Models, Phenomenology and Experimental Results » in 26th International Conference on Supersymmetry and Unification of Fundamental Interactions (SUSY2018), Barcelona, Spain, July 2018.
- Organiser of the session "Higgs Tagging" du workshop ATLAS Boosted Object Tagging (ABOT). April 2016. Heidelberg, Germany.
- Organizer of the session "Jet substructure" of the ATLAS Hadronic Calibration Workshop 2015, September 2015, Bratislava, Slovenia.

B.3 Others

- Member of the comit'e de suivi des th'eses du LPNHE (2021-Present)

- 01/2020-Present, deputy scientific coordinator and CNRS representative for the Erasmus+ Capacity Building project « Latin American alliance for Capacity building in Advanced physics (LA-CoNGA physics) ». Alliance composed of 11 universities from Latin America and Europe, in order to build a virtual learning platform in advanced physics in Colombia, Ecuador, Peru and Venezuela.
- 11/2019-11/2021. Member of the Homeward Bound project 2020-2021. Un programme de leadership, strategy et communication de 12 mois pour des femmes ayant une formation en STEMM. Entre 80 et 100 femmes sont sélectionnées pour suivre cette formation chaque année.
- 04/2019-Present, member of the preparatory group of Latin-American Strategy for Research Infrastructures Group (LASF4RI). LASF4RI is the Latin America equivalent of the European Strategy for Particle Physics.
- Fall semester 2019, Member of the "High Energy" evaluation panel of the National Basic Science Program in the fall semester 2019 for ColCiencias. Colciencias is the administrative department of science, technology and innovation for the government of Colombia. Promotes public policies for the promotion of STI in Colombia.
- 2016-2019, Co-coordinator of the International Centre for Theoretical Physics (ICTP) Physics without Frontiers for Latin America. The team has organized projects in many countries around the world. Personally I have planned and directed activities in Venezuela, Colombia, Peru Peru, Ecuador, Uruguay and Argentina.
- 09/2014-09/2016, Co-director of the CEVALE2VE virtual organization

Bibliography

[1] In: ().

[2] S. Agostinelli et al. "GEANT4 – a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[3] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[4] Leandro G. Almeida et al. "Top Jets at the LHC". In: *Phys. Rev.* D 79 (2009), p. 074012. DOI: 10.1103/PhysRevD.79.074012. arXiv: 0810.0934 [hep-ph].

[5] Guido Altarelli, B. Mele, and M. Ruiz-Altaba. "Searching for New Heavy Vector Bosons in $p\bar{p}$ Colliders". In: *Z. Phys.* C45 (1989). [Erratum: Z. Phys.C47,676(1990)], p. 109. DOI: 10.1007/BF01552335,10.1007/BF01556677.

[6] Lukas Altenkamp et al. "Gluon-induced Higgs-strahlung at next-to-leading order QCD". In: *JHEP* 02 (2013), p. 078. DOI: 10.1007/JHEP02(2013)078. arXiv: 1211.5015 [hep-ph].

[7] J. R. Andersen et al. "Discovering Technicolor". In: *Eur. Phys. J. Plus* 126 (2011), p. 81. DOI: 10.1140/epjp/i2011-11081-1. arXiv: 1104.1255 [hep-ph].

[8] J. R. Andersen et al. "Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report". In: *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*. 2016. arXiv: 1605.04692 [hep-ph].

[9] ATLAS Collaboration. "A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery". In: *Nature* 607.7917 (2022), pp. 52–59. DOI: 10.1038/s41586-022-04893-w. arXiv: 2207.00092 [hep-ex].

[10] ATLAS Collaboration. *A Search for Resonances Decaying to a W or Z Boson and a Higgs Boson in the $q\bar{q}^{(\prime)}b\bar{b}$ Final State*. ATLAS-CONF-2016-083. 2016. URL: https://cds.cern.ch/record/2206276.

[11] ATLAS Collaboration. "A search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1} of pp collisions at $\sqrt{s}$ = 7 TeV". In: *submitted to Phys. Rev.* D (2013). arXiv: 1305.2756 [hep-ex].

[12] ATLAS Collaboration. *A study of Standard Model Higgs boson production in the decay mode $H \to b\bar{b}$ in association with a W or Z boson for High Luminosity LHC Running*. ATL-PHYS-PUB-2014-011. 2014. URL: https://cds.cern.ch/record/1740962.

[13] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: (2019). arXiv: 1907.05120 [hep-ex].

[14] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report*. ATLAS-TDR-19. 2010. URL: https://cds.cern.ch/record/1291633.

[15] ATLAS Collaboration. "ATLAS measurements of the properties of jets for boosted particle searches". In: *Phys. Rev. D* 86 (2012), p. 072006. DOI: 10.1103/PhysRevD.86.072006. arXiv: 1206.5369 [hep-ex].

[16] ATLAS Collaboration. *ATLAS Sensitivity to the Standard Model Higgs in the HW and HZ Channels at High Transverse Momenta*. ATL-PHYS-PUB-2009-088. 2009. URL: https://cds.cern.ch/record/1201444.

[17] ATLAS Collaboration. *Boosted Higgs ($\rightarrow b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATLAS-CONF-2016-039. 2016. URL: https://cds.cern.ch/record/2206038.

[18] ATLAS Collaboration. *Combination of measurements of Higgs boson production in association with a W or Z boson in the $b\bar{b}$ decay channel with the ATLAS experiment at $\sqrt{s} = 13$ TeV*. ATLAS-CONF-2021-051. 2021. URL: http://cds.cern.ch/record/2782535.

[19] ATLAS Collaboration. "Combination of searches for heavy resonances decaying into bosonic and leptonic final states using 36 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 98 (2018), p. 052008. DOI: 10.1103/PhysRevD.98.052008. arXiv: 1808.02380 [hep-ex].

[20] ATLAS Collaboration. *Combination of searches for heavy resonances using 139 fb^{-1} of proton–proton collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. ATLAS-CONF-2022-028 », year =.

[21] ATLAS Collaboration. "Combination of searches for WW, WZ, and ZZ resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Lett. B* 755 (2016), p. 285. DOI: 10.1016/j.physletb.2016.02.015. arXiv: 1512.05099 [hep-ex].

[22] ATLAS Collaboration. *Combined measurements of Higgs boson production and decay using up to 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*. Geneva, 2021. URL: https://cds.cern.ch/record/2789544.

[23] ATLAS Collaboration. *Constituent-Based Top-Quark Tagging with the ATLAS Detector*. Geneva, 2022. URL: https://cds.cern.ch/record/2825328.

[24] ATLAS Collaboration. *Constituent-level pile-up mitigation techniques in ATLAS*. ATLAS-CONF-2017-065. 2017. URL: https://cds.cern.ch/record/2281055.

[25] ATLAS Collaboration. "Constraints on Higgs boson production with large transverse momentum using $H \rightarrow b\bar{b}$ decays in the ATLAS detector". In: *Phys. Rev. D* 105 (2022), p. 092003. DOI: 10.1103/PhysRevD.105.092003. arXiv: 2111.08340 [hep-ex].

[26] ATLAS Collaboration. *Cross-section measurements of the Higgs boson decaying to a pair of tau leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*. ATLAS-CONF-2018-021. 2018. URL: https://cds.cern.ch/record/2621794.

[27] ATLAS Collaboration. "Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector". In: *Eur. Phys. J. C* 80.12 (2020), p. 1104. DOI: 10.1140/epjc/s10052-020-08477-8. arXiv: 1910.04482 [hep-ex].

[28] ATLAS Collaboration. "Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector". In: *Eur. Phys. J. C* 82 (2022), p. 717. DOI: 10.1140/epjc/s10052-022-10588-3. arXiv: 2201.11428 [hep-ex].

[29] ATLAS Collaboration. *Early Inner Detector Tracking Performance in the 2015 data at sqrt(s) = 13 TeV*. ATL-PHYS-PUB-2015-051. 2015. URL: http://cdsweb.cern.ch/record/2110140.

[30] ATLAS Collaboration. *Efficiency corrections for a tagger for boosted $H \rightarrow b\bar{b}$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Geneva, 2021. URL: https://cds.cern.ch/record/2777811.

[31] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data". In: *JINST* 14 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[32] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *JINST* 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[33] ATLAS Collaboration. *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*. ATL-PHYS-PUB-2017-022. 2017. URL: https://cds.cern.ch/record/2298955.

[34] ATLAS Collaboration. "Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector". In: *JHEP* 12 (2017), p. 024. DOI: 10.1007/JHEP12(2017)024. arXiv: 1708.03299 [hep-ex].

[35] ATLAS Collaboration. *Expected Performance of Boosted Higgs ($\to b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-035. 2015. URL: https://cds.cern.ch/record/2042155.

[36] ATLAS Collaboration. *Expected performance of the ATLAS b-tagging algorithms in Run-2*. ATL-PHYS-PUB-2015-022. 2015. URL: https://cds.cern.ch/record/2037697.

[37] ATLAS Collaboration. *Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC*. Geneva, 2021. URL: https://cds.cern.ch/record/2776651.

[38] ATLAS Collaboration. *Flavor Tagging with Track-Jets in Boosted Topologies with the ATLAS Detector*. ATL-PHYS-PUB-2014-013. 2014. URL: https://cds.cern.ch/record/1750681.

[39] ATLAS Collaboration. *Identification of Boosted, Hadronically-Decaying W and Z Bosons in $\sqrt{s} = 13$ TeV Monte Carlo Simulations for ATLAS*. ATL-PHYS-PUB-2015-033. 2015. URL: https://cds.cern.ch/record/2041461.

[40] ATLAS Collaboration. "Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J. C* 76 (2016), p. 154. DOI: 10.1140/epjc/s10052-016-3978-z. arXiv: 1510.05821 [hep-ex].

[41] ATLAS Collaboration. "Identification of boosted Higgs bosons decaying into b-quark pairs with the ATLAS detector at 13 TeV". In: *Eur. Phys. J. C* 79.10 (2019), p. 836. DOI: 10.1140/epjc/s10052-019-7335-x. arXiv: 1906.11005 [hep-ex].

[42] ATLAS Collaboration. *Identification of Boosted Higgs Bosons Decaying Into $b\bar{b}$ With Neural Networks and Variable Radius Subjets in ATLAS*. Geneva, 2020. URL: https://cds.cern.ch/record/2724739.

[43] ATLAS Collaboration. *Identification of hadronically-decaying top quarks using UFO jets with ATLAS in Run 2*. All figures including auxiliary figures are available at https://atlas.web.cern.ch/A PHYS-PUB-2021-028. Geneva, 2021. URL: http://cds.cern.ch/record/2776782.

[44] ATLAS Collaboration. *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. ATL-PHYS-PUB-2017-003. 2017. URL: https://cds.cern.ch/record/2255226.

[45] ATLAS Collaboration. *Impact of Alternative Inputs and Grooming Methods on Large-R Jet Reconstruction in ATLAS*. ATL-PHYS-PUB-2017-020. 2017. URL: https://cds.cern.ch/record/2297485.

[46] ATLAS Collaboration. "Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 73 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex].

[47] ATLAS Collaboration. *Improving jet substructure performance in ATLAS using Track-CaloClusters*. ATL-PHYS-PUB-2017-015. 2017. URL: https://cds.cern.ch/record/2275636.

[48] ATLAS Collaboration. "In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector". In: *Eur. Phys. J. C* 79 (2019), p. 135. DOI: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex].

[49] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-015. 2015. URL: https://cds.cern.ch/record/2037613.

[50] ATLAS Collaboration. "Jet energy measurement with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 7$ TeV". In: *Eur. Phys. J. C* 73 (2013), p. 2304. DOI: 10.1140/epjc/s10052-013-2304-2. arXiv: 1112.6426 [hep-ex].

[51] ATLAS Collaboration. "Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (July 2020). arXiv: 2007.02645 [hep-ex].

[52] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[53] ATLAS Collaboration. *Large-R (anti-kt R=1.0) JES and JMS for 2012 data*. http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-001/.

[54] ATLAS Collaboration. *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. ATLAS-CONF-2019-021. 2019. URL: https://cds.cern.ch/record/2677054.

[55] ATLAS Collaboration. "Luminosity determination in *pp* collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 76 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[56] ATLAS Collaboration. "Measurement of Higgs boson decay into *b*-quarks in associated production with a top-quark pair in *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 06 (2022), p. 097. DOI: 10.1007/JHEP06(2022)097. arXiv: 2111.06712 [hep-ex].

[57] ATLAS Collaboration. "Measurement of k_T splitting scales in $W \to \ell\nu$ events at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 73 (2013), p. 2432. DOI: 10.1140/epjc/s10052-013-2432-8. arXiv: 1302.1415 [hep-ex].

[58] ATLAS Collaboration. "Measurement of the associated production of a Higgs boson decaying into *b*-quarks with a vector boson at high transverse momentum in *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 816 (2021), p. 136204. DOI: 10.1016/j.physletb.2021.136204. arXiv: 2008.02508 [hep-ex].

[59] ATLAS Collaboration. *Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s} = 13$ TeV pp data*. ATLAS-CONF-2020-022. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONF-2020-022. 2020. URL: http://cds.cern.ch/record/2724442.

[60] ATLAS Collaboration. "Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector". In: *Eur. Phys. J. C* 82.1 (2022), p. 95. DOI: 10.1140/epjc/s10052-021-09843-w. arXiv: 2109.10627 [hep-ex].

[61] ATLAS Collaboration. "Measurement of the jet mass in high transverse momentum $Z(\to b\bar{b})\gamma$ production at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: (2019). arXiv: 1907.07093 [hep-ex].

[62] ATLAS Collaboration. "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *Eur. Phys. J. C* 79 (2019), p. 205. DOI: 10.1140/epjc/s10052-019-6650-6. arXiv: 1810.05087 [hep-ex].

[63] ATLAS Collaboration. "Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 121 (2018), p. 092001. DOI: 10.1103/PhysRevLett.121.092001. arXiv: 1711.08341 [hep-ex].

[64] ATLAS Collaboration. "Measurement of VH, $H \to b\bar{b}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector". In: *JHEP* 05 (2019), p. 141. DOI: 10.1007/JHEP05(2019)141. arXiv: 1903.04618 [hep-ex].

[65] ATLAS Collaboration. "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV". In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].

[66] ATLAS Collaboration. *Measurements of Higgs boson production cross-sections in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. ATLAS-CONF-2021-044. 2021. URL: https://cds.cern.ch/record/2779179.

[67] ATLAS Collaboration. "Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 81.6 (2021), p. 537. DOI: 10.1140/epjc/s10052-021-09192-8. arXiv: 2011.08280 [hep-ex].

[68] ATLAS Collaboration. "Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81.2 (2021), p. 178. DOI: 10.1140/epjc/s10052-020-08677-2. arXiv: 2007.02873 [hep-ex].

[69] ATLAS Collaboration. *Methodology for EFT interpretation of Higgs boson Simplified Template Cross-section results in ATLAS*. ATL-PHYS-PUB-2019-042. 2019. URL: https://cds.cern.ch/record/2694284.

[70] ATLAS Collaboration. *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. URL: https://cds.cern.ch/record/2044941.

[71] ATLAS Collaboration. *Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration*. Geneva, 2020. URL: https://cds.cern.ch/record/2718610.

[72] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[73] ATLAS Collaboration. *New techniques for jet energy scale calibration in ATLAS for Run-2*. Geneva, 2022. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2022-004/.

[74] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[75] ATLAS Collaboration. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Phys. Lett. B* 786 (2018), p. 59. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[76] ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. ATL-PHYS-PUB-2017-013. 2017. URL: https://cds.cern.ch/record/2273281.

[77] ATLAS Collaboration. *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. ATL-PHYS-PUB-2016-012. 2016. URL: https://cds.cern.ch/record/2160731.

[78] ATLAS Collaboration. *Optimisation of the smoothing of b-jet identification efficiency and mistag rate simulation-to-data scale factors in ATLAS*. Geneva, 2020. URL: https://cds.cern.ch/record/2710598.

[79] ATLAS Collaboration. "Performance of b-jet identification in the ATLAS experiment". In: *JINST* 11 (2016), P04008. DOI: 10.1088/1748-0221/11/04/P04008. arXiv: 1512.01094 [hep-ex].

[80] ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J. C* 80.1 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[81] ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *JHEP* 09 (2013), p. 076. DOI: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].

[82] ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *JHEP* 09 (2013), p. 076. DOI: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].

[83] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction in proton–proton collisions at $\sqrt{s} = 7$ TeV with ATLAS". In: *Eur. Phys. J. C* 72 (2012), p. 1844. DOI: 10.1140/epjc/s10052-011-1844-6. arXiv: 1108.5602 [hep-ex].

[84] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 78 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[85] ATLAS Collaboration. "Performance of the ATLAS trigger system in 2015". In: *Eur. Phys. J. C* 77 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[86] ATLAS Collaboration. "Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC". In: *Eur. Phys. J. C* 79 (2019), p. 375. DOI: 10.1140/epjc/s10052-019-6847-8. arXiv: 1808.07858 [hep-ex].

[87] ATLAS Collaboration. *Performance of W/Z taggers using UFO jets in ATLAS*. ATL-PHYS-PUB-2021-029. 2021. URL: https://cds.cern.ch/record/2777009.

[88] ATLAS Collaboration. "Search for a new heavy gauge-boson resonance decaying into a lepton and missing transverse momentum in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS experiment". In: *Eur. Phys. J. C* 78 (2018), p. 401. DOI: 10.1140/epjc/s10052-018-5877-y. arXiv: 1706.04786 [hep-ex].

[89] ATLAS Collaboration. "Search for chargino and neutralino production in final states with a Higgs boson and missing transverse momentum at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* (2018). arXiv: 1812.09432 [hep-ex].

[90] 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". In: *JHEP* 09 (2019). [Erratum: JHEP 06, 042 (2020)], p. 091. DOI: 10.1007/JHEP09(2019)091. arXiv: 1906.08589 [hep-ex].

[91] 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". In: (2019). arXiv: 1906.08589 [hep-ex].

[92] ATLAS Collaboration. "Search for diboson resonances with boson-tagged jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 777 (2018), p. 91. DOI: 10.1016/j.physletb.2017.12.011. arXiv: 1708.04445 [hep-ex].

[93] ATLAS Collaboration. "Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector". In: *JHEP* 03 (2018), p. 174. DOI: 10.1007/JHEP03(2018)174. arXiv: 1712.06518 [hep-ex].

[94] ATLAS Collaboration. "Search for heavy resonances decaying into WW in the $e\nu\mu\nu$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78 (2018), p. 24. DOI: 10.1140/epjc/s10052-017-5491-4. arXiv: 1710.01123 [hep-ex].

[95] ATLAS Collaboration. *Search for Heavy Resonances Decaying to a W or Z Boson and a Higgs Boson in the $q\bar{q}^{(\prime)}b\bar{b}$ Final State in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*. ATLAS-CONF-2017-018. 2017. URL: https://cds.cern.ch/record/2258132.

[96] ATLAS Collaboration. "Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the $q\bar{q}^{(\prime)}b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 774 (2017), p. 494. DOI: 10.1016/j.physletb.2017.09.066. arXiv: 1707.06958 [hep-ex].

[97] ATLAS Collaboration. "Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78 (2018), p. 293. DOI: 10.1140/epjc/s10052-018-5686-3. arXiv: 1712.06386 [hep-ex].

[98] ATLAS Collaboration. "Search for Higgs boson production in association with a high-energy photon via vector-boson fusion with decay into bottom quark pairs at $\sqrt{s}$=13 TeV with the ATLAS detector". In: *JHEP* 03 (2021), p. 268. DOI: 10.1007/JHEP03(2021)268. arXiv: 2010.13651 [hep-ex].

[99] ATLAS Collaboration. "Search for high-mass diboson resonances with boson-tagged jets in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 12 (2015), p. 055. DOI: 10.1007/JHEP12(2015)055. arXiv: 1506.00962 [hep-ex].

[100] ATLAS Collaboration. "Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 12 (2015), p. 055. DOI: 10.1007/JHEP12(2015)055. arXiv: 1506.00962 [hep-ex].

[101] ATLAS Collaboration. "Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 10 (2017), p. 182. DOI: 10.1007/JHEP10(2017)182. arXiv: 1707.02424 [hep-ex].

[102] ATLAS Collaboration. "Search for production of WW/WZ resonances decaying to a lepton, neutrino and jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 75 (2015), p. 209. DOI: 10.1140/epjc/s10052-015-3593-4. arXiv: 1503.04677 [hep-ex].

[103] ATLAS Collaboration. "Search for production of WW/WZ resonances decaying to a lepton, neutrino and jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C75.5 (2015). [Erratum: Eur. Phys. J.C75,370(2015)], p. 209. DOI: 10.1140/epjc/s10052-015-3593-4,10.1140/epjc/s10052-015-3425-6. arXiv: 1503.04677 [hep-ex].

[104] ATLAS Collaboration. *Search for resonances with boson-tagged jets in 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*. ATLAS-CONF-2015-073. 2015. URL: https://cds.cern.ch/record/2114845.

[105] ATLAS Collaboration. "Search for resonant diboson production in the $\ell\ell q\bar{q}$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 75 (2015), p. 69. DOI: 10.1140/epjc/s10052-015-3261-8. arXiv: 1409.6190 [hep-ex].

[106] ATLAS Collaboration. "Search for resonant diboson production in the $\ell\ell q\bar{q}$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J.* C75 (2015), p. 69. DOI: 10.1140/epjc/s10052-015-3261-8. arXiv: 1409.6190 [hep-ex].

[107] ATLAS Collaboration. "Search for resonant diboson production in the WW/WZ$\rightarrow$ $\ell\nu jj$ decay channels with the ATLAS detector at $\sqrt{s}$ = 7 TeV". In: *Phys. Rev.* D87.11 (2013), p. 112006. DOI: 10.1103/PhysRevD.87.112006. arXiv: 1305.0125 [hep-ex].

[108] ATLAS Collaboration. "Search for resonant WZ production in the fully leptonic final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 787 (2018), p. 68. DOI: 10.1016/j.physletb.2018.10.021. arXiv: 1806.01532 [hep-ex].

[109] ATLAS Collaboration. "Search for WW/WZ resonance production in $\ell\nu qq$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 03 (2018), p. 042. DOI: 10.1007/JHEP03(2018)042. arXiv: 1710.07235 [hep-ex].

[110] ATLAS Collaboration. "Search for WZ resonances in the fully leptonic channel using pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector". In: *Phys. Lett.* B737 (2014), pp. 223–243. DOI: 10.1016/j.physletb.2014.08.039. arXiv: 1406.4456 [hep-ex].

[111] ATLAS Collaboration. "Searches for heavy diboson resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 09 (2016), p. 173. DOI: 10.1007/JHEP09(2016)173. arXiv: 1606.04833 [hep-ex].

[112] ATLAS Collaboration. "Searches for heavy ZZ and ZW resonances in the $\ell\ell qq$ and $\nu\nu qq$ final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 03 (2018), p. 009. DOI: 10.1007/JHEP03(2018)009. arXiv: 1708.09638 [hep-ex].

[113] ATLAS Collaboration. *Secondary vertex finding for jet flavour identification with the ATLAS detector*. ATL-PHYS-PUB-2017-011. 2017. URL: https://cds.cern.ch/record/2270366.

[114] ATLAS Collaboration. *Studies of b-tagging performance and jet substructure in a high p_T $g \rightarrow b\bar{b}$ rich sample of large-R jets from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*. ATLAS-CONF-2016-002. 2016. URL: https://cds.cern.ch/record/2135187.

[115] ATLAS Collaboration. "Summary of ATLAS Pythia 8 tunes". In: ATL-PHYS-PUB-2012-003,
http://cdsweb.cern.ch/record/1474107 (2012).

[116] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[117] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[118] ATLAS Collaboration. "The performance of the jet trigger for the ATLAS detector during 2011 data taking". In: *Eur. Phys. J. C* 76 (2016), p. 526. DOI: 10.1140/epjc/s10052-016-4325-0. arXiv: 1606.07759 [hep-ex].

[119] ATLAS Collaboration. *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. ATL-PHYS-PUB-2018-025. 2018. URL: https://cds.cern.ch/record/2645405.

[120] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[121] ATLAS Collaboration. *Trigger menu in 2018*. Geneva, 2019. URL: https://cds.cern.ch/record/2693402.

[122] ATLAS Collaboration. *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\rightarrow b\bar{b}$) Tagging in ATLAS*. ATL-PHYS-PUB-2017-010. 2017. URL: https://cds.cern.ch/record/2268678.

[123] *ATLAS Luminosity Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[124] G et al. Avoni. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13 (2018), P07017. 33 p. DOI: 10.1088/1748-0221/13/07/P07017. URL: http://cds.cern.ch/record/2633501.

[125] V. Barger, W. Y. Keung, and Ernest Ma. "Gauge model with light W and Z bosons". In: *Phys. Rev. D* 22 (3 1980), pp. 727–737. DOI: 10.1103/PhysRevD.22.727. URL: https://link.aps.org/doi/10.1103/PhysRevD.22.727.

[126] Alexander Belyaev, Neil D. Christensen, and Alexander Pukhov. "CalcHEP 3.4 for collider physics within and beyond the Standard Model". In: *Comput.Phys.Commun.* 184 (2013), pp. 1729–1769. DOI: 10.1016/j.cpc.2013.01.014. arXiv: 1207.6082 [hep-ph].

[127] Peter Berta et al. "Particle-level pileup subtraction for jets and jet shapes". In: *JHEP* 06 (2014), p. 092. DOI: 10.1007/JHEP06(2014)092. arXiv: 1403.3108 [hep-ex].

[128] Daniele Bertolini et al. "Pileup Per Particle Identification". In: *JHEP* 10 (2014), p. 59. DOI: 10.1007/JHEP10(2014)059. arXiv: 1407.6013 [hep-ph].

[129] R. Blair et al. *The CDF-II detector: Technical design report*. FERMILAB-DESIGN-1996-01. 1996. URL: https://cds.cern.ch/record/1478626.

[130] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders". In: *Phys. Lett. B* 579 (2004), pp. 149–156. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206.

[131] Oliver Brein, Robert V. Harlander, and Tom J.E. Zirke. "vh@nnlo—Higgs Strahlung at hadron colliders". In: *Computer Physics Communications* 184.3 (2013), pp. 998–1003. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2012.11.002. URL: https://www.sciencedirect.com/science/article/pii/S0010465512003785.

[132] Oliver Brein et al. "Top-quark mediated effects in hadronic Higgs-Strahlung". In: *Eur. Phys. J. C* 72 (2012), p. 1868. DOI: 10.1140/epjc/s10052-012-1868-6. arXiv: 1111.0761 [hep-ph].

[133] Ilaria Brivio, Yun Jiang, and Michael Trott. "The SMEFTsim package, theory and tools". In: *JHEP* 12 (2017), p. 070. DOI: 10.1007/JHEP12(2017)070. arXiv: 1709.06492 [hep-ph].

[134] Jonathan M. Butterworth et al. "Jet Substructure as a New Higgs Search Channel at the Large Hadron Collider". In: *Phys. Rev. Lett.* 100 (2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. arXiv: 0802.2470 [hep-ph].

[135] M. Cacciari and G. P. Salam. "Pileup subtraction using jet areas". In: *Phys. Lett.* B 659 (2008), p. 119. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[136] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "SoftKiller, a particle-level pileup removal method". In: *Eur. Phys. J.* C 75 (2015), p. 59. DOI: 10.1140/epjc/s10052-015-3267-2. arXiv: 1407.0408 [hep-ph].

[137] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[138] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Catchment Area of Jets". In: *JHEP* 04 (2008), p. 005. DOI: 10.1088/1126-6708/2008/04/005. arXiv: 0802.1188 [hep-ph].

[139] Reina Coromoto Camacho Toro, Giordon Stark, and David Miller. "gFEX, the ATLAS Calorimeter Global Feature Extractor for the Phase-I upgrade of the ATLAS experiment". In: *PoS* ICHEP2016 (2017), p. 1055. DOI: 10.22323/1.282.1055.

[140] John M. Campbell, R. Keith Ellis, and Ciaran Williams. "Associated production of a Higgs boson at NNLO". In: *JHEP* 06 (2016), p. 179. DOI: 10.1007/JHEP06(2016)179. arXiv: 1601.00658 [hep-ph].

[141] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[142] S. Catani et al. "Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions". In: *Nucl. Phys. B* 406 (1993), pp. 187–224. DOI: 10.1016/0550-3213(93)90166-M.

[143] S. Catani et al. "Longitudinally invariant $k_\perp$ clustering algorithms for hadron hadron collisions". In: *Nucl. Phys.* B 406 (1993), p. 187. DOI: 10.1016/0550-3213(93)90166-M.

[144] Simon Catterall et al. "MCRG Minimal Walking Technicolor". In: *Phys. Rev.* D85 (2012), p. 094501. DOI: 10.1103/PhysRevD.85.094501. arXiv: 1108.3794 [hep-ph].

[145] M. Cepeda et al. "Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC". In: *CERN Yellow Rep. Monogr.* 7 (2019). Ed. by Andrea Dainese et al., pp. 221–584. DOI: 10.23731/CYRM-2019-007.221. arXiv: 1902.00134 [hep-ph].

[146] Chunhui Chen. "New approach to identifying boosted hadronically-decaying particle using jet substructure in its center-of-mass frame". In: *Phys. Rev.* D 85 (2012), p. 034007. DOI: 10.1103/PhysRevD.85.034007. arXiv: 1112.2567 [hep-ph].

[147] M. L. Ciccolini, S. Dittmaier, and M. Krämer. "Electroweak radiative corrections to associated WH and ZH production at hadron colliders". In: *Phys. Rev. D* 68 (2003), p. 073003. DOI: 10.1103/PhysRevD.68.073003. arXiv: hep-ph/0306234.

[148] Xabier Cid Vidal et al. "Report from Working Group 3: Beyond the Standard Model physics at the HL-LHC and HE-LHC". In: *CERN Yellow Rep. Monogr.* 7 (2019). Ed. by Andrea Dainese et al., pp. 585–865. DOI: 10.23731/CYRM-2019-007.585. arXiv: 1812.07831 [hep-ph].

[149] CMS Collaboration. *Direct search for the standard model Higgs boson decaying to a charm quark-antiquark pair*. 2022. URL: https://cds.cern.ch/record/2802742.

[150] CMS Collaboration. "Identification techniques for highly boosted W bosons that decay into hadrons". In: *JHEP* 12 (2014), p. 017. DOI: 10.1007/JHEP12(2014)017. arXiv: 1410.4227 [hep-ex].

[151] CMS Collaboration. "Inclusive search for highly boosted Higgs bosons decaying to bottom quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 12 (2020), p. 085. DOI: 10.1007/JHEP12(2020)085. arXiv: 2006.13251 [hep-ex].

[152] CMS Collaboration. "Measurements of the differential production cross sections for a Z boson in association with jets in pp collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 04 (2017), p. 022. DOI: 10.1007/JHEP04(2017)022. arXiv: 1611.03844 [hep-ex].

[153] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* 716 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[154] CMS Collaboration. "Observation of Higgs Boson Decay to Bottom Quarks". In: *Phys. Rev. Lett.* 121 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[155] CMS Collaboration. "Observation of the Higgs boson decay to a pair of τ leptons". In: *Phys. Lett. B* 779 (2018), p. 283. DOI: 10.1016/j.physletb.2018.02.004. arXiv: 1708.00373 [hep-ex].

[156] CMS Collaboration. "Search for $\mathrm{t\bar{t}H}$ production in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel with leptonic $\mathrm{t\bar{t}}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 03 (2019), p. 026. DOI: 10.1007/JHEP03(2019)026. arXiv: 1804.03682 [hep-ex].

[157] CMS Collaboration. "Search for exotic resonances decaying into WZ/ZZ in pp collisions at $\sqrt{s} = 7$ TeV". In: *JHEP* 02 (2013), p. 036. DOI: 10.1007/JHEP02(2013)036. arXiv: 1211.5779 [hep-ex].

[158] CMS Collaboration. "Search for heavy resonances in the W/Z-tagged dijet mass spectrum in pp collisions at 7 TeV". In: *Phys. Lett.* B723 (2013), pp. 280–301. DOI: 10.1016/j.physletb.2013.05.040. arXiv: 1212.1910 [hep-ex].

[159] CMS Collaboration. "Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at $\sqrt{s} = 8$ TeV". In: *JHEP* 08 (2014), p. 174. DOI: 10.1007/JHEP08(2014)174. arXiv: 1405.3447 [hep-ex].

[160] CMS Collaboration. "Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 08 (2014), p. 173. DOI: 10.1007/JHEP08(2014)173. arXiv: 1405.1994 [hep-ex].

[161] CMS Collaboration. "Search for massive WH resonances decaying into the $\ell\nu b\bar{b}$ final state at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J. C* 76 (2016), p. 237. DOI: 10.1140/epjc/s10052-016-4067-z. arXiv: 1601.06431 [hep-ex].

[162] CMS Collaboration. "Studies of jet mass in dijet and W/Z + jet events". In: *JHEP* 05 (2013), p. 090. DOI: 10.1007/JHEP05(2013)090. arXiv: 1303.4811 [hep-ex].

[163] ATLAS Collaboration. "A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV". In: *Eur. Phys. J. C* 72 (2012), p. 2083. DOI: 10.1140/epjc/s10052-012-2083-1. arXiv: 1205.5371 [hep-ex].

[164] ATLAS Collaboration. "Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions". In: *Eur. Phys. J. C* 81.4 (2021), p. 334. DOI: 10.1140/epjc/s10052-021-09054-3. arXiv: 2009.04986 [hep-ex].

[165] CMS Collaboration. *Combined Higgs boson production and decay measurements with up to 137 fb-1 of proton-proton collision data at sqrts = 13 TeV*. Geneva, 2020. URL: https://cds.cern.ch/record/2706103.

[166] CMS Collaboration. *Simplified template cross section measurements of Higgs boson produced in association with vector bosons in the H→b$\bar{b}$ decay channel in proton-proton collisions at* $\sqrt{s} = 13$ TeV. Geneva, 2022. URL: https://cds.cern.ch/record/2827421.

[167] CMS Collaboration. *W and top tagging scale factors for Run 2 data*. 2020. URL: https://cds.cern.ch/record/2718978.

[168] Roberto Contino et al. "Effective Lagrangian for a light Higgs-like scalar". In: *JHEP* 07 (2013), p. 035. DOI: 10.1007/JHEP07(2013)035. arXiv: 1303.3876 [hep-ph].

[169] Roberto Contino et al. "On the effect of resonances in composite Higgs phenomenology". In: *Journal of High Energy Physics* 2011.10 (2011). DOI: 10.1007/jhep10(2011)081. URL: https://doi.org/10.1007%2Fjhep10%282011%29081.

[170] Chris Malena Delitzsch. "Search for Diboson Resonances in the Fully Hadronic Final State Using Jet Substructure Techniques in 8 and 13 TeV Proton-Proton Collisions with the ATLAS Detector". Presented 02 Dec 2016. 2016. URL: https://cds.cern.ch/record/2242565.

[171] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph].

[172] Frédéric A. Dreyer and Huilin Qu. "Jet tagging in the Lund plane with graph networks". In: *JHEP* 03 (2021), p. 052. DOI: 10.1007/JHEP03(2021)052. arXiv: 2012.08526 [hep-ph].

[173] Frédéric A. Dreyer, Gavin P. Salam, and Grégory Soyez. "The Lund Jet Plane". In: *JHEP* 12 (2018), p. 064. DOI: 10.1007/JHEP12(2018)064. arXiv: 1807.04758 [hep-ph].

[174] Estia Eichten and Kenneth Lane. "Low-scale technicolor at the Tevatron and LHC". In: *Phys. Lett.* B669 (2008), pp. 235–238. DOI: 10.1016/j.physletb.2008.09.047. arXiv: 0706.2339 [hep-ph].

[175] John Ellis, Veronica Sanz, and Tevong You. "Complete Higgs Sector Constraints on Dimension-6 Operators". In: *JHEP* 07 (2014), p. 036. DOI: 10.1007/JHEP07(2014)036. arXiv: 1404.3667 [hep-ph].

[176] S. D. Ellis, C. K. Vermilion, and J. R. Walsh. "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches". In: *Phys. Rev.* D 81 (2010), p. 094023. DOI: 10.1103/PhysRevD.81.094023. arXiv: 0912.0033 [hep-ph].

[177] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Phys. Rev.* D 48 (1993), p. 3160. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266 [hep-ph].

[178] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088/1748-0221/3/08/s08001.

[179] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano. "Associated Higgs-*W*-Boson Production at Hadron Colliders: a Fully Exclusive QCD Calculation at NNLO". In: *Phys. Rev. Lett.* 107 (2011), p. 152003. DOI: 10.1103/PhysRevLett.107.152003. arXiv: 1107.1164 [hep-ph].

[180] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano. "Associated ZH production at hadron colliders: The fully differential NNLO QCD calculation". In: *Phys. Lett. B* 740 (2015), pp. 51–55. DOI: 10.1016/j.physletb.2014.11.040. arXiv: 1407.4747 [hep-ph].

[181] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano. "Higher-order QCD effects for associated WH production and decay at the LHC". In: *JHEP* 04 (2014), p. 039. DOI: 10.1007/JHEP04(2014)039. arXiv: 1312.1669 [hep-ph].

[182] Geoffrey C. Fox and Stephen Wolfram. "Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes". In: *Phys. Rev. Lett.* 41 (1978), p. 1581. DOI: 10.1103/PhysRevLett.41.1581.

[183] Harald Fritzsch and Peter Minkowski. "Unified Interactions of Leptons and Hadrons". In: *Annals Phys.* 93 (1975), pp. 193–266. DOI: 10.1016/0003-4916(75)90211-0.

[184] H. Georgi and S. L. Glashow. "Unity of All Elementary Particle Forces". In: *Phys. Rev. Lett.* 32 (1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438.

[185] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[186] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[187] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[188] B. Grzadkowski et al. "Dimension-Six Terms in the Standard Model Lagrangian". In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[189] Robert V. Harlander, Stefan Liebler, and Tom Zirke. "Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-Doublet Model". In: *JHEP* 02 (2014), p. 023. DOI: 10.1007/JHEP02(2014)023. arXiv: 1307.8122 [hep-ph].

[190] Robert V. Harlander et al. "Soft gluon resummation for gluon-induced Higgs Strahlung". In: *JHEP* 11 (2014), p. 082. DOI: 10.1007/JHEP11(2014)082. arXiv: 1410.0217 [hep-ph].

[191] B. Hespel, F. Maltoni, and E. Vryonidou. "Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM". In: *JHEP* 06 (2015), p. 065. DOI: 10.1007/JHEP06(2015)065. arXiv: 1503.01656 [hep-ph].

[192] D. Krohn, J. Thaler, and L.-T. Wang. "Jet Trimming". In: *JHEP* 02 (2010), p. 084. DOI: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[193] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jets with Variable R". In: *JHEP* 06 (2009), p. 059. DOI: 10.1088/1126-6708/2009/06/059. arXiv: 0903.0392 [hep-ph].

[194] Andrew J. Larkoski, Ian Moult, and Duff Neill. "Analytic Boosted Boson Discrimination". In: (2015). arXiv: 1507.03018 [hep-ph].

[195] Andrew J. Larkoski, Ian Mount, and Duff Neill. "Power Counting to Better Jet Observables". In: *JHEP* 12 (2014), p. 009. DOI: 10.1007/JHEP12(2014)009. arXiv: 1409.6298 [hep-ph].

[196] Andrew J. Larkoski, Duff Neill, and Jesse Thaler. "Jet Shapes with the Broadening Axis". In: *JHEP* 04 (2014), p. 017. DOI: 10.1007/JHEP04(2014)017. arXiv: 1401.2158 [hep-ph].

[197] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. "Energy Correlation Functions for Jet Substructure". In: *JHEP* 06 (2013), p. 108. DOI: 10.1007/JHEP06(2013)108. arXiv: 1305.0007 [hep-ph].

[198] Andrew J. Larkoski et al. "Soft drop". In: *JHEP* 05 (2014), p. 146. DOI: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph].

[199] LHC Higgs Cross Section Working Group. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". In: *CERN-2013-004* (2013). DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].

[200] LHC Higgs Cross Section Working Group. "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector". In: *CERN-2017-002* (2017). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[201] *Longer term LHC schedule*. URL: https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm.

[202] A.D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J. C* 63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[203] Ken Mimasu, Veronica Sanz, and Ciaran Williams. "Higher Order QCD predictions for Associated Higgs production with anomalous couplings to gauge bosons". In: *JHEP* 08 (2016), p. 039. DOI: 10.1007/JHEP08(2016)039. arXiv: 1512.02572 [hep-ph].

[204] L. Moneta et al. "The RooStats project". In: *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research. February 22-27*. 2010.

[205] Pavel M. Nadolsky et al. "Implications of CTEQ global analysis for collider observables". In: *Phys. Rev.* D 78 (2008), p. 013004. DOI: 10.1103/PhysRevD.78.013004. arXiv: 0802.0007 [hep-ph].

[206] Duccio Pappadopulo et al. "Heavy Vector Triplets: Bridging Theory and Data". In: *JHEP* 09 (2014), p. 060. DOI: 10.1007/JHEP09(2014)060. arXiv: 1402.4431 [hep-ph].

[207] B. Pifer et al. "An Experiment at D0 to Study anti-Proton - Proton Collisions at 2-TeV: Design Report". In: FERMILAB-PUB-83-111-E. PRINT-84-0306 (FERMILAB). FERMILAB-DESIGN-1983-04 (1983). URL: https://cds.cern.ch/record/1478625.

[208] Huilin Qu and Loukas Gouskos. "ParticleNet: Jet Tagging via Particle Clouds". In: *Phys. Rev. D* 101.5 (2020), p. 056019. DOI: 10.1103/PhysRevD.101.056019. arXiv: 1902.08570 [hep-ph].

[209] Lisa Randall and Raman Sundrum. "A Large mass hierarchy from a small extra dimension". In: *Phys. Rev. Lett.* 83 (1999), pp. 3370–3373. DOI: 10.1103/PhysRevLett.83.3370. arXiv: hep-ph/9905221 [hep-ph].

[210] Lisa Randall and Raman Sundrum. "An Alternative to compactification". In: *Phys. Rev. Lett.* 83 (1999), pp. 4690–4693. DOI: 10.1103/PhysRevLett.83.4690. arXiv: hep-th/9906064 [hep-th].

[211] Alexander L. Read. "Presentation of search results: The CL(s) technique". In: *J.Phys.* G28 (2002), pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[212] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Proceedings of the 8th Nobel symposium, Ed. N. Svartholm, Almqvist & Wiskell, 1968,* Conf. Proc. C680519 (1968), pp. 367–377. URL: {http://inspirehep.net/record/53083}.

[213] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[214] J. Thaler and L.-T. Wang. "Strategies to Identify Boosted Tops". In: *JHEP* 07 (2008), p. 092. DOI: 10.1088/1126-6708/2008/07/092. arXiv: 0806.0023 [hep-ph].

[215] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. DOI: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph].

[216] Wouter Verkerke and David Kirkby. *The RooFit toolkit for data modeling*. 2003. arXiv: physics/0306116 [physics.data-an].

[217] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[218] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.