Instytut Fizyki Jądrowej
im. Henryka Niewodniczańskiego
Polskiej Akademii Nauk
ul. Radzikowskiego 152, 31-342 Kraków

www.ifj.edu.pl/badania/publikacje

Kraków, March 2019

Selected measurements of final states with third-generation fermions in proton – proton collisions using the ATLAS detector at the LHC

Paweł Malecki

Habilitation Thesis

Wydano Nakładem Instytutu Fizyki Jądrowej im. Henryka Niewodniczańskiego
Polskiej Akademii Nauk
Kraków 2019

Recenzent: dr hab. Paweł Brückman de Renstrom

ISBN 978-83-63542-09-2

Streszczenie

Niniejsza monografia koncentruje się na temacie pomiarów prowadzonych w ostatnich latach w eksperymencie ATLAS na Wielkim Zderzaczu Hadronów z wykorzystaniem stanów końcowych zawierających fermiony trzeciej generacji. Dyskutowane wyniki dotyczą zderzeń proton – proton przy energiach w środku masy wynoszących $\sqrt{s} = 7$, 8 i 13 TeV, zarejestrowanych podczas tzw. Run 1 oraz pierwszej połowy Run 2.

Algorytmy używane w eksperymencie ATLAS do rekonstrukcji i identyfikacji hadronowych rozpadów leptonów τ a także dżetów hadronowych zawierających kwark b są szczegółowo opisane, zarówno w wersjach używanych w Run 1 jak i w Run 2, gdzie zaimplementowano szereg poprawek i usprawnień. Prezentowane są także wyniki pomiarów wydajności tych algorytmów oraz częstości mylnej identyfikacji nieprawdziwych leptonów τ lub dżetów b.

Monografia przedstawia wybrane pomiary (i wynikające z ich analizy wnioski) z obszaru Modelu Standardowego, włączając pomiar polaryzacji leptonów τ pochodzących z rozpadu bozonu Z, a także pomiary i wnioski dotyczące fizyki bozonu Higgsa z Modelu Standardowego, rozpadającego się na parę $\tau\tau$ lub bb. W kolejnych rozdziałach prezentowane są pomiary mające na celu poszukiwanie dodatkowych bozonów Higgsa spoza Modelu Standardowego, włączając ciężki elektrycznie obojętny bozon Higgsa poszukiwany w rozpadzie na dwa leptony τ oraz na dwa kwarki b. Ponadto, opisywane są także poszukiwania naładowanego bozonu Higgsa rozpadającego się na stany końcowe zawierające pary $\tau\nu$ lub tb. W żadnej z opisywanych analiz nie znaleziono znaczących odchyleń od przewidywań Modelu Standardowego, zostały jednak oszacowane i zaprezentowane limity wykluczeń.

Abstract

This monograph is focused on the measurements performed by the ATLAS experiment using final states containing third-generation fermions in the recent years and covers both the Run-1 and early Run-2 data taking periods of proton – proton collisions at $\sqrt{s} = 7$, 8 and 13 TeV centre-of-mass energies.

The ATLAS algorithms for reconstructing and identifying the hadronically-decaying τ leptons and hadronic jets containing b-hadrons are described in detail, covering both the end-of-Run-1 status and Run-2 developments. In addition, experimental measurements of the efficiencies of these algorithms are presented together with the results on their misidentification rates in background objects.

Selected measurements (and derived conclusions) in the Standard Model sector are presented, including the measurement of τ polarisation in Z-boson decays and measurements in the SM Higgs sector with both $\tau\tau$ and bb final states. This is followed by a presentation of selected searches for non-Standard-Model Higgs bosons, including the heavy neutral one in di-τ and di-b decays as well as the searches for the charged Higgs boson in $\tau\nu$ and tb channels. No deviation from Standard-Model predictions is observed in any of the presented analyses and no hint of non-SM Higgs boson existence was found. Exclusion limits are presented instead.

Author's contribution

The processes with third-generation fermions have been my research interest since my joining the ATLAS collaboration in 2006 and started with τ leptons.

Following my PhD defence in 2011, I started working on the algorithms for reconstruction and identification of the hadronic decays of τ leptons in the ATLAS experiment. I was coordinating the work on the development and optimisation of the TauDiscriminant software package containing the cut-based and multivariate methods for τ identification, that was used throughout the whole ATLAS collaboration in all analyses with hadronic τ decays in the final states. I was also the main developer of the cluster-based Pi0Finder algorithm, used to reconstruct and count neutral pions in hadronic τ decays in the Run-1 data. I also took part in the preparation of the Run-2 τ reconstruction algorithm, the TauParticleFlow, especially in the topics related to optimisation of the π^0 reconstruction. I therefore contributed to many parts of the work described in Section 4.1. I was not involved in the part on b-jet reconstruction and tagging described in Section 4.2. The presented techniques and results on efficiency and misidentification rates measurements are described for completeness of this monograph.

Meanwhile, in 2012 I started studies on the polarisation of τ leptons in $Z/\gamma^* \to \tau\tau$ decays. I made the first design of the analysis strategy (optimised event selection criteria, method for extracting the actual polarisation), and was not only taking active part in its further development, statistical analysis and studies on systematic uncertainties, but was also coordinating it as analysis leader throughout its whole course, including being the publication editor. Therefore, my role in the analysis described in Section 5.1 was the central one.

I was not directly involved in the Run-1 $H \to \tau\tau$ search described in Section 5.2, however my work related to τ-identification algorithms had a direct impact on that search that would not be possible without robust τ identification methods. I did not take part in the analysis of Standard-Model Higgs decays to $b\bar{b}$, described in Section 5.3, however it fully coincides with the subject of this monograph and is therefore presented.

Following the ATLAS physics workshop at Aix-les-Bains in 2014 my physics interest in the Beyond-Standard-Model Higgs sector had started. I joined a group that was planning to conduct a search for heavy neutral Higgs boson decaying into $b\bar{b}$ pair and produced in association with at least one b quark. During the course of that analysis I became a principal investigator, contributing to all stages of the analysis, including studies on simulation samples, optimisation of event selection and background estimation. My main contribution was however the construction of the fit model and performing the whole statistical analysis including analysis of systematic uncertainties. I played a key role in that search, described in Section 6.1, including being one of the publication editors. The BSM $H \to \tau\tau$ search is presented in this monograph for completeness (Section 6.2) as a complementary result to the previously mentioned one. I was, however, not involved in it.

The searches for charged Higgs bosons are also described in this monograph to present a more complete picture of the status of experimental results in the BSM Higgs sector. I did not take part in the $H^+ \to tb$ search, described in Section 6.4. I was a co-supervisor of a PhD student who took active part in the $H^+ \to \tau\nu$ search, described in Section 6.3. Her developments in the area of multivariate analysis were important for this search, whereas the studies on the use of embedding technique yielded promising results, but due to time

constraints were not included in the current version of the analysis published by ATLAS, but are considered for the full-Run-2 version of this search.

Following the rules of the ATLAS collaboration, only public plots and results are presented in this monograph.

Contents

Chapter 1

Introduction

In the early days of the Standard Model (SM) two generation of elementary fermions seemed sufficient to describe the Universe [1–3]. First-generation quarks (up and down) together with electrons form almost all of the known baryonic matter. Observations of muons in cosmic rays as well as the *strange* behaviour of kaons that were produced in strong interactions but were able to decay only via weak processes made the second generation necessary. There was neither any experimental hint nor a theoretical need for the third generation. However, a CP-violation mechanism proposed in 1973 by Kobayashi and Maskawa [4] made the hypothesis of its existence a serious possibility to consider.

Shortly after, in 1975, a first experimental claim for τ lepton observation was published [5], followed by later observations in 1978 that made its existence established [6]. Meanwhile, in 1977, the E288 experiment in Fermilab observed the *bottomonium*, a state composed of $b\bar{b}$ quark-anti-quark pair [7]. Since then, the properties of both these fermions were thoroughly studied by many experiments, such as Belle [8, 9], BaBar [10, 11], the LEP experiments [12, 13] and many others.

Currently, in the age of high-energy hadron colliders (Tevatron [14], Large Hadron Collider, LHC [15]) both the τ leptons and b quarks are used as probes for properties of their production processes such as decays of electroweak bosons, top quarks, Higgs bosons and possible New Physics (NP) particles.

The τ is the heaviest ($m_\tau = 1.776$ GeV [16]) lepton, the only one that can decay into hadrons. Its short (2.9×10^{-13} s [16]) lifetime causes these decays to happen within detector volume making the reconstruction of their products possible. Because of that the τ is well-suited for measurements of polarisation, spin correlations and CP numbers of the decaying resonances, especially as the decay modes of τ lepton are precisely known from low-energy experiments. The reconstruction of hadronic τ decays is however very challenging at hadron colliders due to overwhelming background from QCD multi-jet production that makes efficient selection of true τ lepton decays very difficult. Also, efficient triggering of events with hadronic τ decays with low-enough trigger rates forms another challenge.

The bottom quarks, next-to-heaviest ($m_b = 4.18$ GeV [16]) fundamental fermions, decay into lighter quarks via weak interactions at rates determined by the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements [4]. Resulting long (10^{-12} s) lifetimes of mesons containing the b quark allow for efficient distinction of hadronic jets formed by them from lighter-quark and gluon jets by means of so-called *b-tagging*. These methods exploits information about displaced vertices in collision events, decay multiplicities and b mass. The high masses of bottom quarks makes the $H \to b\bar{b}$ decay the most frequent

($\mathcal{B} \approx 58\%$) decay channel of the SM Higgs boson. Also, in many extensions of the Standard Model, such as Two-Higgs-Doublet Models (2HDM) [17, 18] the decays of the possible heavy objects into b-quark pairs could be very important. This makes final states with b jets an excellent place to examine the properties of the Higgs boson and to search for New Physics phenomena.

On the other hand, the top (t) quark was only discovered in 1995 [19, 20]. Its properties (mass, branching fractions, etc.) are still themselves subject to interest and are studied in the LHC experiments [21, 22]. The large mass (173 GeV [16]) of the top quark makes it less-frequently used as a probe for underlying processes.

The tau neutrino, ν_τ, the last element of the third generation, the existence of which became imminent after the discovery of τ lepton, was first directly observed by the DONUT experiment [23] in 2000. In collider experiments it is not directly measurable but its presence in collision events, as well as of other neutrino flavours, can only be detected by observing non-zero total transverse momentum. This however may form an important signature for many processes involving decays of charged objects such as $W^\pm$ or Beyond-Standard-Model charged Higgs bosons.

The LHC has been in operation since 2009 and has been colliding protons at centre-of-mass energies of $\sqrt{s} = 7$, 8 (so-called Run 1) and 13 TeV (Run 2) allowing the ATLAS experiment [24] to collect 5.1 fb^{-1}, 21.3 fb^{-1} and 149 fb^{-1} with these three energies respectively[1]. This monograph presents a selection of results from Run 1 and from a subset of data collected in Run 2 (up to 2017). No public results from the presented topics were available for the full Run-2 dataset analyses by the time of writing of this monograph.

The document is organised as follows. Chapter 2 gives a short description of properties of τ leptons and bottom quarks and reviews processes with those fermions in final states. In Chapter 3 a description of the experimental apparatus, the ATLAS detector, is given. Subsystems are briefly described as well as techniques for particle identification. Chapter 4 describes the algorithms for τ reconstruction and identification as well as the b tagging. Chapters 5 and 6 describe the measurements and searches in the Standard Model Sector and Beyond Standard Model (BSM), respectively.

Conventions and notation

The system of units used assumes the speed of light $c = 1$. This implies that both the momentum and mass are expressed in units of energy, eV (usually: GeV or TeV).
Electric charge is expressed in terms of elementary charge so proton has charge = 1.
Charge-conjugate states are implied unless stated otherwise, so i.e. charged Higgs boson is denoted H^+ and H^- is implied, the same holds for other particles.
In many cases electric charge is omitted, so, i.e. τ means τ^+ or τ^-.
Hadronically-decaying τ is marked with τ_{had} whereas a leptonically-decaying one is written as τ_{lep}.
Neutrino flavours are usually omitted (but are clear from the context).
Electrons and muons are commonly denoted as leptons and marked ℓ. If the τ leptons are considered this is stated explicitly.
In the Standard-Model context, the Higgs boson is marked as H. The H is also used for the BSM heavy neutral CP-even Higgs boson but the distinction should be clear from the context.
The most-frequently used acronyms are listed in Table 1.1.

[1] 1 fb^{-1} corresponds approximately to 10^{14} collisions at the LHC.

Acronym	Meaning
2HDM	Two-Higgs-Doubled Models - one of the simplest extensions to the SM introducing extended Higgs sector with five bosons instead of one
BDT	Boosted Decision Tree - a multivariate analysis algorithm
BSM	Beyond Standard Model
DM	Dark Matter
EF	Event Filter, the highest level of the ATLAS trigger system
EM	Electromagnetic (usually when describing parts of calorimeter)
ggF	Gluon-gluon fusion, a Higgs-boson production mechanism
HLT	High-Level Trigger, containing Level 2 (L2) and Event Filter (EF)
ID	Inner Detector of the ATLAS experiment, responsible for tracking measurements
JER	Jet energy resolution
JES	Jet energy scale
L1	Level-1 trigger
L2	Level-2 trigger
LCW	Local Cell Weighting, a technique for correcting the energy of hadrons measured in the ATLAS calorimeter system
ML	Machine Learning
MS	Muon Spectrometer of the ATLAS detector
MSSM	Minimal Supersymmetric Standard Model - a minimal extension to the Standard Model that realises supersymmetry
MVA	Multivariate Analysis
NP	New Physics (equivalent to Beyond Standard Model, BSM)
OS	Opposite Signs (of electric charges)
QCD	Quantum Chromodynamics, a theory governing strong interactions of quarks and gluons
SCT	Semi-conductor Tracker, a tracking sub-detector of the ATLAS Inner Detector
SM	Standard Model of elementary particles
SS	Same Signs (of electric charges)
SUSY	Supersymmetry, one of many theories that could extend the Standard Model
TER	Tau energy resolution
TES	Tau energy scale
TRT	Transition Radiation Tracker - a gaseous tracking detector with electron/hadron identification capabilities, part of the ATLAS Inner Detector (ID)
VBF	Vector Boson Fusion, a Higgs-boson production mechanism
VH	Vector-associated Higgs-boson production, also denoted Higgsstrahlung

Table 1.1: Frequently-used acronyms.

Chapter 2

Physics with τ leptons and b quarks

The third generation of fermions is composed of the τ and τ-neutrino in the lepton sector and of the top (t) and bottom (b) quarks. In the era of LHC and hadron-collider experimental environment the b quarks and τ leptons are primarily used to examine the nature of their production processes including both the Standard-Model measurements and New Physics searches. The properties of top quarks are still being examined (also with τ- and b-final states) whereas the neutrinos are studied in other, non-collider experiments that are outside of the scope of this monograph.

2.1 Physics with τ leptons

The τ lepton was the first observed member of the third generation. Its properties have been studied for decades in many experiments and are well established allowing to use it as a probe for SM and non-SM processes. In this Section, in addition to the properties of τ leptons, also the basic aspects of the production of the SM and non-SM Higgs bosons are introduced. This information is relevant also for the subsequent Section.

2.1.1 Properties of τ leptons

The τ lepton mass is 1776.86 ± 0.12 MeV and its lifetime is $(2.903 \pm 0.005) \times 10^{-13}$ s with corresponding decay length of 87.03 μm [16]. Thus, unlike electrons and muons, in LHC experiments the τ's decay inside detectors, not even reaching their innermost layers.

The τ^- couples to a W^- boson producing a τ neutrino. The virtual W^- boson couples further to either a pair of leptons ($e^-\bar{\nu}_e$ or $\mu^-\bar{\nu}_\mu$) or quarks ($\bar{u}d$ or $\bar{u}s$, as other quark pairs such as $\bar{c}d$ and $\bar{c}s$ would be too heavy). This results in two main decay channels of the τ lepton, the leptonic and hadronic one, respectively:

- Leptonic decays (τ_{lep}):
  $\tau^- \to \nu_\tau \bar{\nu}_e e^-$
  $\tau^- \to \nu_\tau \bar{\nu}_\mu \mu^-$
- Hadronic decays (τ_{had}):
  $\tau^- \to \nu_\tau \bar{u} d$
  $\tau^- \to \nu_\tau \bar{u} s$

All of the τ decays, to lowest order, are contained in the list presented above. Couplings to quark pairs are determined by the CKM matrix elements [4] that strongly favour the

pairs from the same-generation ($\bar{u}d$) over the cross-generation pairs ($\bar{u}s$) resulting in much larger branching fraction for the former of the hadronic decays. Figure 2.1 illustrates the decay channels of τ^- lepton listed above.

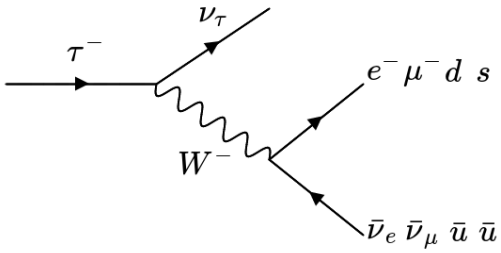


Figure 2.1: Illustration of τ^- decay channels.

The most important branching fractions are given in Table 2.1. The quark pairs produced in W^- decays may further produce additional quark-anti-quark pairs ($u\bar{u}$, $d\bar{d}$ or $s\bar{s}$) resulting in a number of mesons (π, K) forming final states of the process. The leptonic

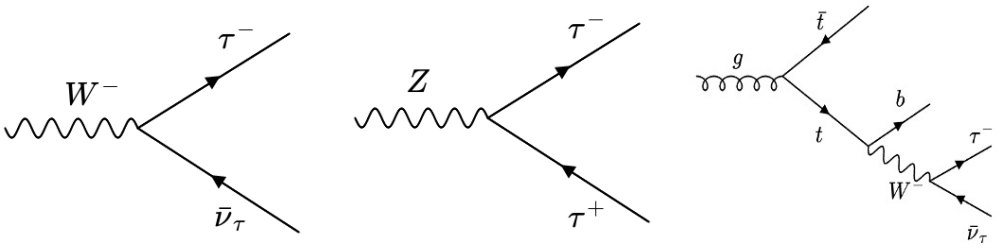
W^- decay	τ decays hadron channel	final state	Branching fraction [%]
$e^-\bar{\nu}_e$		e^-	17.82 ± 0.04
$\mu^-\bar{\nu}_\mu$		μ^-	17.39 ± 0.04
$\bar{u}d$	π^-	π^-	10.82 ± 0.05
$\bar{u}d$	ρ^-	$\pi^-\pi^0$	25.49 ± 0.09
$\bar{u}d$	a_1^-	$\pi^-\pi^+\pi^-$	8.99 ± 0.05
$\bar{u}d$	a_1^-	$\pi^-\pi^0\pi^0$	9.26 ± 0.10
$\bar{u}d$		$\pi^-\pi^+\pi^-\pi^0$	4.62 ± 0.05
$\bar{u}d$		$\pi^-\pi^0\pi^0\pi^0$	1.04 ± 0.07
$\bar{u}s$	K^-	K^-	0.70 ± 0.01
$\bar{u}s$	K^{*-}	$K^-\pi^0$	0.43 ± 0.02
$\bar{u}s$	K^{*-}	K^-K^0	0.15 ± 0.01

Table 2.1: Most frequent τ^- decay modes. The decays are grouped according to the W^- decay channels. Branching fraction values are current world average [16].

decays of τ leptons occur in 35% of the cases whereas hadronic decays are almost twice more frequent and occur in 65% of the cases. Out of the hadronic decays, 77% contain a single charged meson and are called *single-prong* (*1-prong*, *1p*), 23% have three charged mesons (denoted *three-prong*, *3-prong*, *3p*) and only 0.1% contain 5 of more *prongs*. In the final states the $\pi^\pm$ and π^0 mesons are dominant but small fractions of both charged and neutral kaons are also present.

2.1.2 Standard Model processes with τ leptons

In the LHC Run 1 the measurements of SM processes with τ leptons in final states were of great importance for the development of experimental methods of τ detection. They helped calibrating the reconstructed τ lepton energies and helped to understand the performance of τ reconstruction and identification algorithms. Moreover, an important part in the ATLAS physics programme was to measure and understand the previously known processes at energies much higher than available before as these processes form important backgrounds for Higgs and New-Physics searches. The two dominant sources of τ leptons in proton – proton collisions at the LHC are the decays of weak bosons, $W \to \tau\nu$ and $Z \to \tau\tau$. Also, the decays of top quarks, especially in $t\bar{t}$ production were subject to study.



(a) (b) (c)

Figure 2.2: Main Standard Model production channels of τ leptons

The $W \to \tau\nu$ process (Fig. 2.2(a)) is the most abundant source of τ's in pp collisions at the LHC. Its cross-section times branching ratio is $\sigma \times \mathcal{B} = 10.46 \pm 0.52$ nb at $\sqrt{s} = 7$ TeV [25–28]. This process, in addition to good τ reconstruction, requires a robust measurement of transverse missing energy that appears in such events due to neutrinos escaping detection. The $Z/\gamma^* \to \tau\tau$ decay (Fig. 2.2(b)) delivers two τ leptons in the final state, yet its cross-section times branching ratio is 10 times smaller than that of $W \to \tau\nu$ [25–27]. Nevertheless, the presence of two τ's with near-Z invariant mass makes it a very good channel to study the performance of reconstruction, identification and triggering of hadronic decays of τ leptons, especially in the τ_{lep}-τ_{had} final state with so-called *tag-and-probe* approach, where the leptonic *leg* is used for selecting the events and the hadronic *leg* can be used for the aforementioned studies. More importantly, in both these channels one can measure the cross-section times branching ratio and, consequently, lepton universality by comparing to other leptonic decay channels of W and Z. Moreover, measurements of τ lepton polarisation are possible in these processes. Such measurements had not previously been performed at hadron collider experiments.

Measurement of top quark properties is an important part of the ATLAS physics programme. The $t\bar{t}$ pairs are continuously produced at the LHC. The top quark decays predominantly into bW pair. One of the subsequent decay channels involves τ leptons: $t \to b\tau\nu_\tau$ (Fig. 2.2(c)). It provides an insight into the coupling of third-generation fermions in a single process and is particularly important for the searches for Beyond-Standard-Model charged Higgs bosons as they can be produced via top-quark decays.

2.1.3 Standard Model Higgs boson decays into τ final states

The search for the Higgs boson had been one of the main physics goals of the LHC and the general-purpose detectors, ATLAS and CMS [29]. It was completed by the Higgs-boson discovery in 2012 [30, 31] in $\gamma\gamma$, WW and ZZ decay channels. This opened a new window in experimental physics of the Higgs boson allowing for detailed studies of its production mechanisms, couplings and other properties.

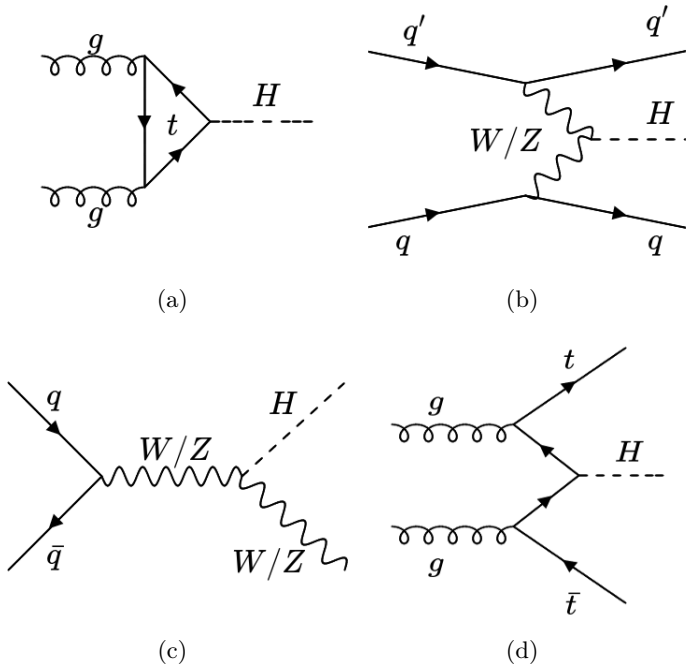


Figure 2.3: Standard Model production of the Higgs boson at the LHC. Top-left: gluon-gluon fusion (ggF), top-right: vector-boson fusion, bottom-left: Higgsstrahlung (VH), bottom-right: top-associated (ttH) production.

At the LHC there are four dominant Higgs production modes [32–34], presented in Fig. 2.3. The most frequent one is the gluon-gluon fusion (ggF) mechanism (Fig. 2.3(a)) [35] that contributes in about 86% to the total Higgs production cross-section. This production channel is experimentally difficult to observe because of lack of clean signature to select ggF events from overwhelming QCD multi-jet background. Currently, only leptonic final states (from Higgs decaying into leptons or W/Z bosons) could probe this mode.

The sub-dominant Higgs-boson production channel is the vector-boson fusion (VBF, Fig. 2.3(b)) [36]. In such processes, in addition to the decay of the Higgs boson, two forward jets can be observed with little or no hadronic activity in the event due to lack of colour connection between the two outgoing quarks. It is also a challenging search channel in collider experiments as the forward jets can escape detector acceptance being too close to the beam pipe.

The VH production mode, often referred to as *Higgsstrahlung* (Fig. 2.3(c)) is experimentally easier to observe because of an additional signature of the process which is a decay of the vector boson accompanying the Higgs, even being only the third most

frequent Higgs-boson production channel.

The fourth, least frequent of the mentioned Higgs production processes, is the top-associated ($t\bar{t}H$) mode (Fig. 2.3(d)). Its contribution to the total Higgs-boson production cross-section is of the order of 1% being already 3-5 times smaller (depending on $\sqrt{s}$) than that of VH.

Even though the first observations of the Higgs particles were performed in their decays into bosons, fermionic decay channels have also been of great importance to LHC experiments. These channels directly probe the Yukawa coupling of the Higgs boson to fermions and may demonstrate the proportionality of coupling strength to the fermion mass. The $H \to \tau\tau$ decay channel was the first fermionic decay channel of the Higgs boson to be observed [37–39], despite the fact that the most frequent one is the $H \to b\bar{b}$ [1] (see Fig. 2.4). The branching ratio of $H \to \tau\tau$ is around 6%. The $H \to \tau\tau$ decays also suffer

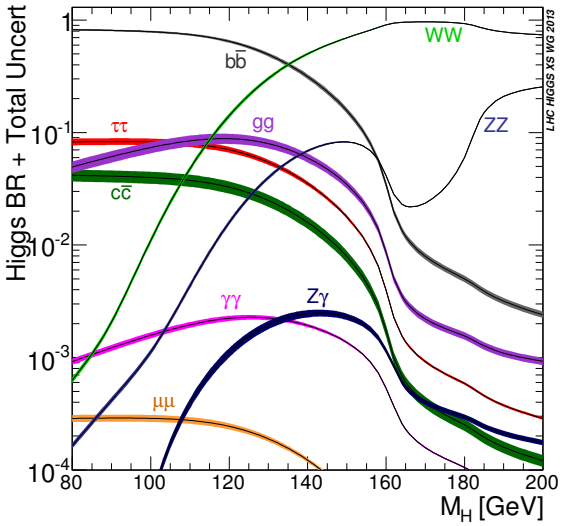


Figure 2.4: Standard Model Higgs boson decay branching ratios, from [34].

from large backgrounds which, however, can be suppressed more easily than in $H \to b\bar{b}$ case by requiring additional jets (VBF-oriented search) or by searching for transversely-boosted Higgs bosons (ggF-oriented search). This is possible because τ reconstruction and identification algorithms provide good rejection of QCD multi-jet background events.

2.1.4 Searches for New Physics with τ leptons

The LHC experiments have a broad programme of searches for Beyond-Standard-Model phenomena in a vast amount of models and scenarios. It would not be possible to cover them all and it is not the purpose of this monograph, so a small subset of considered SM extensions is presented here.

[1] The $H \to b\bar{b}$ decay (described in Chapter 2.2.3) is overwhelmed by enormous backgrounds and requires additional signatures and larger data statistics to be observed.

Additional Higgs bosons

The discovery of the Higgs boson confirms the mechanism of the electroweak symmetry breaking with a doublet of scalar fields [40–42]. Alternative models, containing a scalar particle with SM-like-Higgs properties are not excluded however. A simple extension of the electroweak symmetry breaking mechanism in the SM is the Two-Higgs-Doublet Model (2HDM) [17, 18] containing two doublets of scalar fields instead of one. Each of these fields have vacuum-expectation values and their ratio is usually denoted $\tan\beta$. There are eight degrees of freedom provided by these two fields, of which three are absorbed (as in SM) by longitudinal degrees of freedom of the W and Z bosons and five remain physical and appear as five Higgs bosons, denoted h (neutral light scalar, usually equated with the SM Higgs boson), H (heavy neutral scalar), A (neutral CP-odd particle) and charged $H^{\pm}$'s. The most general 2HDM contains Higgs-mediated flavour-changing neutral currents (FCNS's). They are experimentally strongly limited [18], so selecting appropriate Higgs-to-fermions couplings is needed. In total, four types of the 2HDM emerge that differ between each other by their Yukawa couplings. They are presented in Table 2.2. Out of these types of 2HDM, type-II is realised in supersymmetric (SUSY) [43–47] models such as Minimal-Supersymmetric Standard Model (MSSM) [48–52].

	$h\bar{U}U$ ξ^u_h	$h\bar{D}D$ ξ^d_h	$h\bar{E}E$ ξ^e_h	$H\bar{U}U$ ξ^u_H	$H\bar{D}D$ ξ^d_H	$H\bar{E}E$ ξ^e_H	$iA\bar{U}\gamma_5 U$ ξ^u_A	$iA\bar{D}\gamma_5 D$ ξ^d_A	$iA\bar{E}\gamma_5 E$ ξ^e_A
I	$1+\frac{\epsilon}{t_\beta}$	$1+\frac{\epsilon}{t_\beta}$	$1+\frac{\epsilon}{t_\beta}$	$-(\frac{1}{t_\beta}-\epsilon)$	$-(\frac{1}{t_\beta}-\epsilon)$	$-(\frac{1}{t_\beta}-\epsilon)$	$-\frac{1}{t_\beta}$	$\frac{1}{t_\beta}$	$\frac{1}{t_\beta}$
II	$1+\frac{\epsilon}{t_\beta}$	$1-\epsilon t_\beta$	$1-\epsilon t_\beta$	$-(\frac{1}{t_\beta}-\epsilon)$	$t_\beta+\epsilon$	$t_\beta+\epsilon$	$-\frac{1}{t_\beta}$	$-t_\beta$	$-t_\beta$
X	$1+\frac{\epsilon}{t_\beta}$	$1+\frac{\epsilon}{t_\beta}$	$1-\epsilon t_\beta$	$-(\frac{1}{t_\beta}-\epsilon)$	$-(\frac{1}{t_\beta}-\epsilon)$	$t_\beta+\epsilon$	$-\frac{1}{t_\beta}$	$\frac{1}{t_\beta}$	$-t_\beta$
Y	$1+\frac{\epsilon}{t_\beta}$	$1-\epsilon t_\beta$	$1+\frac{\epsilon}{t_\beta}$	$-(\frac{1}{t_\beta}-\epsilon)$	$t_\beta+\epsilon$	$-(\frac{1}{t_\beta}-\epsilon)$	$-\frac{1}{t_\beta}$	$-t_\beta$	$\frac{1}{t_\beta}$

Table 2.2: Couplings ξ of the 2HDM h, H and A bosons to fermions for 2HDM types I, II, X (or *lepton-specific*) and Y (or *flipped*). The top row explicitly describes the fields involved in the couplings, U, D, and E denote up-type quarks, down-type quarks and charged leptons, respectively, $t_\beta \equiv \tan\beta$ is the ratio of the vacuum expectation values of the two scalar doublets, whereas $\epsilon = \cos(\beta-\alpha)$ where α is the neutral CP-even Higgs mixing angle [17]. The couplings are normalized to the SM values and are given in the alignment limit $\cos(\beta-\alpha) \approx 0$ where the couplings of the light scalar boson h are SM-like. From Ref. [53].

As the Higgs boson, in general, couples to mass, third-generation fermions are natural candidates for final states suitable for 2HDM Higgs-bosons searches. Their dominant coupling is to $b\bar{b}$ (around 90%) but a significant fraction of decays involves τ leptons (around 10%). With increasing values of $\tan\beta$ the couplings can be strongly enhanced following the possible coupling schemes presented in Table 2.2. Therefore, for specific 2HDM types (type-II and lepton-specific), searches with τ final states are best-suited for probing this sector. The $H/A \to \tau\tau$ process can be relevant for a whole range of the heavy Higgs masses up to 1 TeV. Pre-LHC searches in this channel were performed at LEP [54] and at Tevatron [55] and resulted in exclusion of some regions in parameter space.

The Yukawa couplings of the charged Higgs boson are of type:

$$\bar{U}(m_u\xi^u_A + m_d\xi^d_A)D + \bar{\mathcal{V}}m_\ell\xi^\ell_A L \tag{2.1}$$

with $\mathcal{V}$ denoting neutrino and ξ values listed in Table 2.2 [17]. The production process of

the charged Higgs boson, illustrated in Fig. 2.5, depends on the H^+ mass. If it is lower than that of the top quark, the main production channel of charged Higgs bosons is the decay of top quark: $t \to H^+b$ (Fig. 2.5 (a)). For higher H^+ masses, above top-quark threshold, the production mode is via single-resonant top-quark production, $gg \to tbH^+$ (Fig. 2.5 (b)). In the intermediate H^+ mass region the interference between the three main diagrams is also relevant. The dominant decay channel of the charged Higgs boson in the

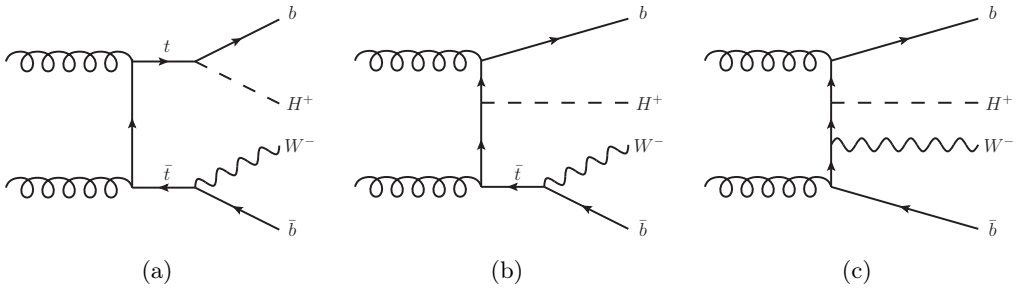


Figure 2.5: Example leading-order Feynman diagrams for the H^+ boson production. Left: double-resonant top-quark production (dominant for low H^+ masses), middle: single-resonant top-quark production (dominant for large H^+ masses), right: non-resonant top-quark production. From Ref. [56].

.

low-mass case (for $\tan\beta > 3$) is $H^+ \to \tau\nu$ [32]. For $m_{H^+} > m_{top}$ the dominant decay mode of the H^+ is $H^+ \to t\bar{b}$ but $H^+ \to \tau\nu$ is also present and remains significant for large $\tan\beta$ values offering experimentally-cleaner observation channel.

Previous searches by the LEP experiments resulted in a lower bound on the charged Higgs mass of 78.6 GeV [57]. The experiments at Tevatron placed upper limits on the $\mathcal{B}(t \to H^+b)$ [58, 59]. Both results were obtained assuming $\mathcal{B}(H^+ \to \tau\nu) = 1$.

Other New Physics searches with τ's

In addition to the type-II of 2HDM, other searches for supersymmetry with τ leptons are proposed, depending on the considered model. In general, in SUSY, the bosonic superpartners of fermions are called sfermions. Their mass eigenstates are mixtures of left- and right-handed components due to electroweak symmetry breaking. In some models large mixing results in that the lightest sfermions belong to the third generation and, together with gauginos (supersymmetric partners of gauge bosons) will decay preferentially to τ leptons. This is for instance realised in the Gauge-Mediated SUSY Breaking (GMSB) scenario [60] where the lightest stau ($\tilde{\tau}$) slepton is the next-to-lightest (NLSP) SUSY particle and the gravitino ($\tilde{G}$) is the lightest SUSY particle, so stau decays into gravitino and tau lepton. Searches for SUSY with τ leptons have been previously performed at LEP, especially at the OPAL experiment, excluding NLSP $\tilde{\tau}$ masses below 87.4 GeV [61]. The D0 experiment at Tevatron was searching for squark production in events with jets, taus and missing transverse energy [62] and, more recently, the ATLAS and CMS collaborations have also published their search results (i.e. Refs. [63, 64]). No significant excess above the SM background was found.

Heavy gauge bosons (W', Z') are high-mass vector-boson resonances proposed by many BSM models such as Sequential Standard Model (SSM) [65] and many others [66–68]. These particles can serve as mediators in Dark Matter (DM) interactions, as proposed by several models [69–71] and are therefore subjects to interest of physicists. In some of them the proposed bosons would couple preferentially to τ leptons [72–74] making τ leptons well-suited to probe these models. Recent results from ATLAS and CMS provide exclusion limits for the SSM excluding the W' up to the mass of 4.0 TeV [75] and Z' up to 2.42 TeV [76].

The decay channels with τ leptons play also an important role in searches for processes with production of two SM Higgs bosons. This broad class of phenomena includes both the SM (e.g. Higgs trilinear couplings) and BSM processes such as decays of heavy resonances. Recently, the ATLAS and CMS collaborations have published results of such searches in $bb\tau\tau$ final states yielding no excess above SM expectations [77, 78].

2.2 Physics with b quarks

The b quark is the next-to-heaviest member of the third generation, constituting a perfect probe for processes involving couplings proportional to mass even despite a more difficult experimental signature with respect to τ leptons and larger experimental backgrounds for such processes.

2.2.1 Properties of b quarks

As proposed by Kobayashi and Maskawa [4], the decays of heavier quarks into lighter ones, and quark-flavour transitions in general are described by the CKM matrix (Eq 2.2). Apart from the fact that it predicts CP-violation in flavour transitions by means of a complex phase appearing in some of its elements, this matrix is nearly diagonal. This means that quark-flavour transitions are strongly favoured within the same generation.

$$\begin{pmatrix} u \\ c \\ t \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{tc} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.2}$$

As the bottom quark is the lighter ($m_b = 4.18 \pm 0.4$ GeV [16]) member of the third-generation quark doublet, it can decay only via generation-changing processes. Their rates are described by the CKM matrix elements $|V_{cb}|$ and $|V_{ub}|$. Both are much smaller ($(42.2 \pm 0.8) \times 10^{-3}$ and $(3.94 \pm 0.36) \times 10^{-3}$, respectively) than one, resulting in relatively long lifetimes (10^{-12} s) of particles containing b quark(s). The most frequent b-quark decay is therefore $b \to cW^{*-}$ where W^{*-} can decay further into lepton-neutrino or quark-anti-quark pairs. The $b \to uW^{*-}$ is suppressed by a factor of $|V_{ub}|^2/|V_{cb}|^2 \approx 0.01$. The b-quark decays are presented in Fig. 2.6.

The hadronic jets initiated by bottom quarks or by hadrons containing bottom quarks display distinctive properties such as displaced secondary production vertices resulting from long lifetime and harder momentum spectra of the daughter particles resulting from high b-quark mass. Around 10% of b jets contain leptons from leptonic W decays. These properties are exploited in experimental identification of b jets.

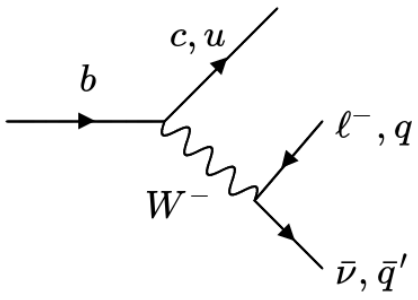


Figure 2.6: b-quark decay modes.

.

2.2.2 b quarks in Standard-Model processes

A significant part of measurements with bottom quarks in the ATLAS experiment is related to top-quark properties. The heaviest third-generation fermion (and - as of now - the heaviest known elementary particle) decays predominantly into a b quark and W boson - the related CKM-matrix element is consistent with one: $|V_{tb}| = 1.019 \pm 0.025$ [16].

At the LHC the top quarks are most frequently produced in top – anti-top ($t\bar{t}$) pairs. A sub-dominant production mode is the single-top quark production. These production modes are presented in Fig. 2.7. In the analyses of top-quark decays many final states are

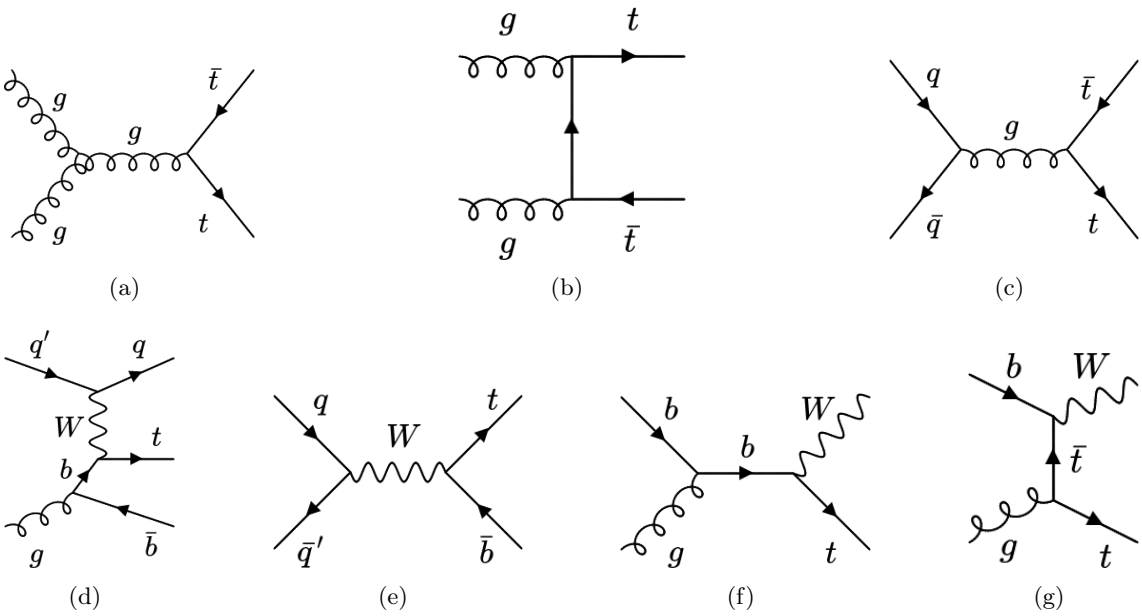


Figure 2.7: Main Standard Model production channels of top-quarks. Top: $t\bar{t}$ pair production, bottom: single-top production.

considered, depending on the W-boson decay channel. However, a single b jet (or a pair

of b jets in $t\bar{t}$ case) is always present in the final state.

2.2.3 b quarks in decays of Standard-Model Higgs boson

As mentioned in Sec. 2.1.3, the discovery of the Higgs boson has allowed for detailed studies of its properties including its couplings to fermions. Even though the $\tau\tau$ decay channel was easier accessible experimentally, it was the $b\bar{b}$ state that was predicted to be most frequent and therefore the measurement of the $H \to b\bar{b}$ branching fraction was expected to provide strongest constraint on the total Higgs-decay width. The searched final state is however overwhelmed by huge background from multi-jet processes, so additional signatures are needed to discriminate against the background. Because of that, the $H \to b\bar{b}$ decays are searched for in the VH production mode that contains an additional vector boson to help selecting signal events. The observation of $H \to b\bar{b}$ decays has been recently reported by both ATLAS [79] and CMS [80] experiments.

2.2.4 Beyond-Standard-Model searches with b quarks

In addition to measurements related to SM phenomena, the b quarks play also an important role in searches for New Physics in LHC experiments. Again, a small subset of all considered searches is presented here.

Additional Higgs bosons

Similarly to $\tau\tau$, the $b\bar{b}$ final state is well-suited for heavy neutral Higgs boson searches. Its couplings are already large at small $\tan\beta$ values (around 90%) and are strongly enhanced at larger $\tan\beta$ for 2HDM of type-II and *flipped*, see Table 2.2. However, similarly to SM case, large experimental backgrounds make such searches challenging and requiring additional signatures to suppress the background. One possibility is to consider b-associated production of heavy Higgs boson where it is produced in association with one or more b quarks. Figure 2.8 presents leading-order diagrams of such processes. The cross-section for such process can be computed in two different so-called flavour schemes. The b-quark mass is much larger than the QCD scale Λ_{QCD} so its production is a perturbative process. In the four-flavour scheme (4FS) the b quark is not considered as a parton in the proton, whereas in the five-flavour scheme (5FS) it is considered massless and its PDF's are taken into account in the cross-section calculation. If all orders of the perturbation series are taken into account, the two approaches are, in fact, identical [81]. The searches for heavy

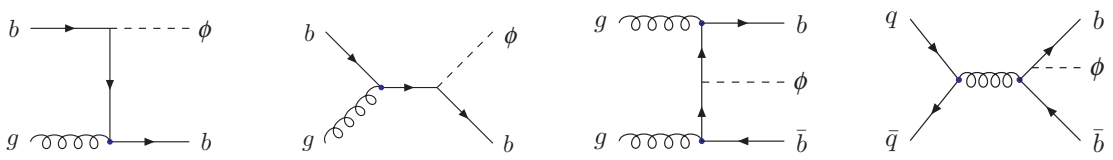


Figure 2.8: Feynman diagrams for some of the leading order processes for the production of a heavy neutral Higgs boson (denoted here by ϕ) in association with one or two b quarks in the 5FS [81].

Higgs boson decaying into bottom-quark pair conducted before the start of the LHC were

performed at LEP [54] and at Tevatron [82] having found no signal and deriving limits in the 90 - 900 GeV mass range.

Other Beyond-Standard-Model phenomena

Bottom quarks play an important role also in the searches for SUSY particles. In simplified models [83, 84] employed by LHC experiments, the gluino (a gluon superpartner) would decay into stop – top or sbottom – bottom pairs with stops and sbottoms decaying further into a neutralino and a top or bottom quark, respectively. The resulting top quark decays further, following the SM, into bW. This in general yields final states with a number of b jets with additional signatures, such as missing transverse energy or leptons. Such searches have been performed at the LHC having found no evidence for searched signal [85–89].

Generic searches for heavy resonances with two b jets in final states are also performed at the LHC. Such resonances with enhanced couplings to $b\bar{b}$ pairs are predicted by some of the Dark Matter models [90, 91]. Both the ATLAS [92] and CMS [93, 94] have not found any evidence for signal existence.

Chapter 3

The ATLAS detector at the Large Hadron Collider

The Large Hadron Collider (LHC) [15] at CERN[1] was designed and constructed to provide collisions of particles at unprecedented energies and with very high intensities. Such conditions are needed for detailed studies on the nature of the Universe, such as precise measurements of SM phenomena but, more importantly, to allow for possible discoveries of previously unobserved processes related to the Higgs sector and New Physics. The LHC has been in operation since 2008 and, following a major damage right after the start, is successfully providing collisions for physics experiments since December 2009. Four main detectors were built to allow for detailed observations of these collisions, one of them being the ATLAS detector [24]. It is the largest of the LHC experiments and is described in more detail later in this Chapter. The measurements described in this monograph were performed using the ATLAS detector.

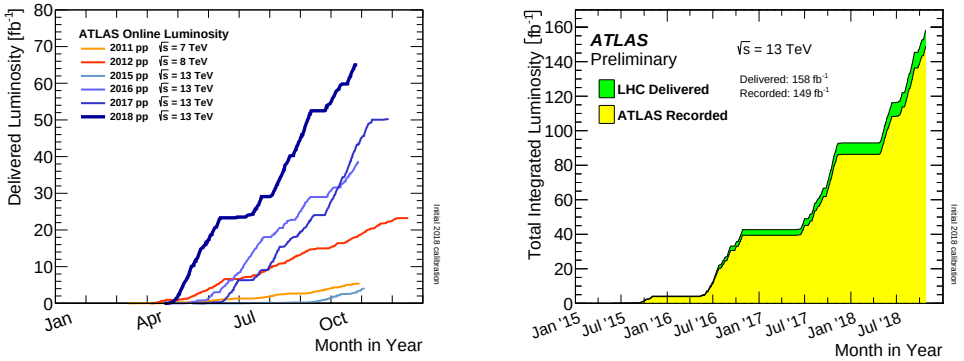
3.1 The Large Hadron Collider

Since its start, the LHC is the particle accelerator with highest energies of accelerated beams in the world. In largest part of its operation time it provides proton – proton collisions but is also used to accelerate and collide heavy ions, such as lead, for a short period every year. The protons are accelerated in groups - so-called bunches, containing about 10^{11} protons, separated in time by 25 ns. Therefore, in each bunch-crossing a number of individual pp collisions can take place. The average number of collisions per bunch crossing ranged from around 6 in 2011 up to 37 in 2018. Usually no more than one such interaction results in a hard process (involving a large momentum transfer compared to the mass of a proton), interesting for further analysis. The other additional interactions are called in-time *pile-up* and contain mostly elastic and inelastic proton – proton scattering, commonly known as Minimum-Bias events. Out-of-time pile-up can also occur, meaning that the remnants of previous bunch-crossings are present in the detector by the time the current bunch-crossing takes place. Large pile-up can affect the performance of event reconstruction and sensitivity of measurements, so special techniques are used for its proper treatment. However, large pile-up is an inevitable consequence of increased instantaneous luminosity that is needed for collecting larger data samples.

[1] European Organization for Nuclear Research, *Conseil Européen pour la Recherche Nucléaire*

In the initial phase, in 2010 and 2011, protons were accelerated up to energies of 3.5 TeV per beam, resulting in collision energies of $\sqrt{s} = 7$ TeV. This period, followed by 2012 collisions at $\sqrt{s} = 8$ TeV are commonly denoted as LHC Run 1. During that period the peak instantaneous luminosity reached $2.0 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$ in 2010, $3.65 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ [95] and $7.7 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ [96] in 2012. The total integrated Run-1 luminosity collected by ATLAS was 28.5 fb^{-1}.

In 2015, following a two-years technical stop, the LHC resumed its operation in the so-called Run 2 period that ended in 2018. The collision energy was raised to $\sqrt{s} = 13$ TeV and a significant increase in instantaneous luminosity was achieved reaching the peak value of $21.4 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ in 2018. The total integrated luminosity recorded by ATLAS in Run 2 was 149 fb^{-1} [97]. The instantaneous luminosity for several data-taking years are summarised in Fig. 3.1 (left) and the total integrated Run-2 luminosity is shown in Fig. 3.1 (right).



(a) (b)

Figure 3.1: Cumulative luminosity versus day delivered to ATLAS during stable beams and for high energy *pp* collisions (left), Cumulative luminosity versus time delivered to ATLAS (green) and recorded by ATLAS (yellow) during stable beams for *pp* collisions at 13 TeV centre-of-mass energy in LHC Run 2 (right) [97].

Four large detectors are constructed at the LHC ring. Two of them are the general-purpose detectors, ATLAS and CMS (Compact Muon Solenoid), built to target a wide variety of physics processes including the searches for non-Standard-Model phenomena. They focus mainly on the physics of proton – proton collisions, but the collisions of heavy ions are also examined in these experiments. The ALICE (A Large Ion Collider Experiment) [98] detector was mainly designed to study heavy-ion collisions at the LHC but is also capable of recording and analysing *pp* events. The LHCb detector [99] on the other hand was designed to examine the physics of heavy flavours in *pp* collisions.

3.2 The ATLAS detector

The ATLAS (A Toroidal LHC ApparatuS) is the largest (44m length, 25m of diameter, 7000 tonnes of weight) of the LHC detectors. It is designed to cover a wide range of particle-physics phenomena, especially searches for Higgs boson and New Physics signatures but also for precise measurements of properties of SM particles and processes. The possible

new particles manifesting Beyond-Standard-Model phenomena can be, in principle, very massive and would possibly decay in many different final states. The detector should therefore be able to efficiently measure momenta, positions and energies of particles produced in each collision, and to provide as much information about the measured objects as possible to allow for robust identification of various particles and, consequently, to help selecting rare and interesting collision events from the overwhelming background of other, mostly multi-jet-production processes. Additionally, its subsystems need to withstand very high radiation and provide fast readout due to very high rates of proton – proton interactions.

The ATLAS is a typical general-purpose particle detector of collider type, consisting of many detector subsystems. All subsystems are geometrically divided in three parts - the central Barrel, with sub-detectors organised in a structure of concentric layers, and two symmetric Endcaps at both ends. The sub-detectors are grouped in such way, that the particle outgoing a collision would first traverse the layers of Inner Detector (ID), that is responsible for tracking. Further behind the electromagnetic and hadronic calorimeters are located. Finally, the outermost layers constitute Muon Spectrometer (MS) system. A schematic overview of the ATLAS detector is presented in Fig. 3.2.

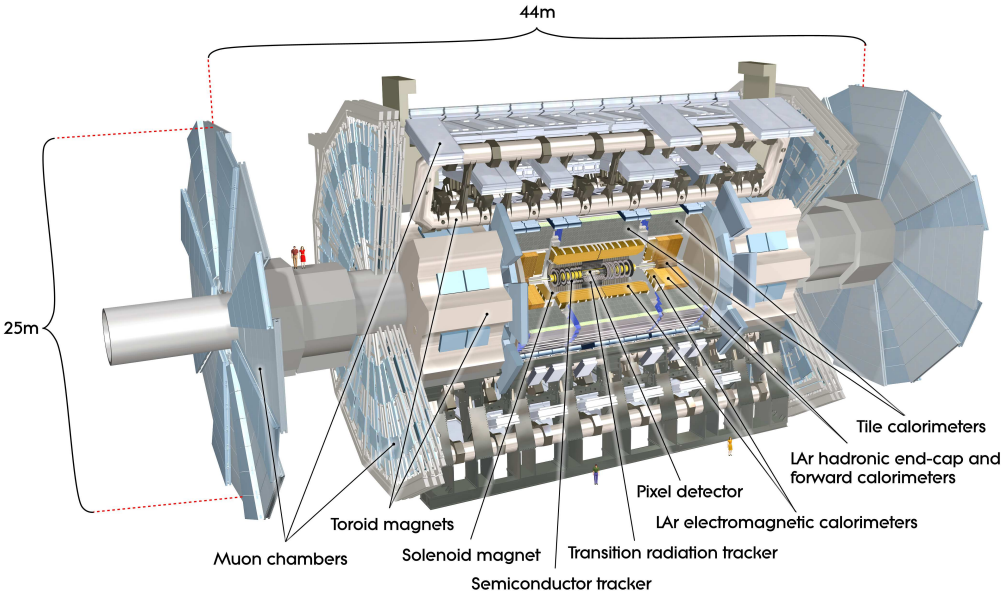


Figure 3.2: Schematic view of the ATLAS detector [24].

3.2.1 Coordinate system and coverage

The ATLAS experiment uses a cylindrical coordinate system. The z-axis is parallel to the beam-line and points towards the anti-clockwise direction of the beam in the LHC ring. The x-axis points towards the centre of the LHC ring and the y-axis points upwards, making the coordinate system right-handed. The azimuthal angle is defined around the beam axis. Its value of $\phi = 0$ corresponds to the direction of the x-axis. The polar angle θ is the angle from the beam-axis. The positive direction of the z-axis corresponds to $\theta = 0$. A commonly used quantity, the pseudorapidity, is defined as $\eta = -\ln\frac{\theta}{2}$. The distance in

$\eta - \phi$ space (usually corresponding to angular separation of tracks, jets, or other objects) is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The transverse quantities used in many ATLAS measurements, such as transverse momentum p_{T}, are defined in the $x - y$ plane.

The ATLAS detector provides an excellent, nearly-4π coverage of the solid angle, with full-2π azimuthal coverage. In the polar angle, or rather in terms of pseudorapidity, the track reconstruction is possible up to $|\eta| < 2.5$, calorimeters cover the area up to $|\eta| < 4.9$ and the accessible range of the forward detectors is $5.4 < |\eta| < 6.1$ and $|\eta| > 8.2$. Muon tracking is possible for $|\eta| < 2.7$. The subsystems of the ATLAS detector are described further and are summarised in Table 3.1 together with their designed performance.

Detector component	Design resolution	η coverage	
		Measurement	Triggering
Tracking	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\% p_{\mathrm{T}} \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry			
Barrel and end-caps	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \|\eta\| < 4.9$	$3.1 < \|\eta\| < 4.9$
Muon spectrometer	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 10\%$ at $p_{\mathrm{T}} = 1$ TeV	± 2.7	± 2.4

Table 3.1: Designed performance and coverage of the ATLAS detector subsystems. The energy and p_{T} are expressed in GeV [24].

3.2.2 Inner Detector

The reconstruction of tracks of charged particles is possible thanks to the Inner Detector, the innermost subsystem of the ATLAS setup. It provides precise measurements of positions of particles along their flight path. The ID is immersed in a $B = 2T$ solenoid magnetic field parallel to the z-axis, therefore providing the possibility of transverse-momentum and charge-sign determination. Three main parts of the Inner Detector can be singled out. The ID is depicted in Fig. 3.3.

Pixel detector

In Run 1, the Pixel detector consisted of three layers of silicon pixel sensors organised in three concentric cylinders in the barrel and three disks perpendicular to the beam axis in each endcap. High granularity of the sensors is required to provide robust measurements in a very dense environment, especially to efficiently reconstruct primary and secondary vertices and to discriminate against pile-up collisions. The sensors in these three layers are organised in such way that the position resolution in the $R\phi$ direction is 10 μm and 115 μm in the z (or R in endcaps) direction. In total, the Pixel system had 80M readout channels. Typically, a charged particle traversing the Pixel system produces 3 hits in the detector.

Before the start of Run 2 a significant upgrade of the Pixel detector was made by adding an additional, innermost layer of pixels named Insertable B-Layer (IBL). The IBL was inserted 3.3 cm from the beam axis, between the previously-innermost layer of pixels (5 cm from the beam axis) and the beam pipe. It consists of 6M of sensors and its resolution in $R\phi$ direction is similar to that of the rest of the Pixel detector whereas the z-resolution is 60 μm [100].

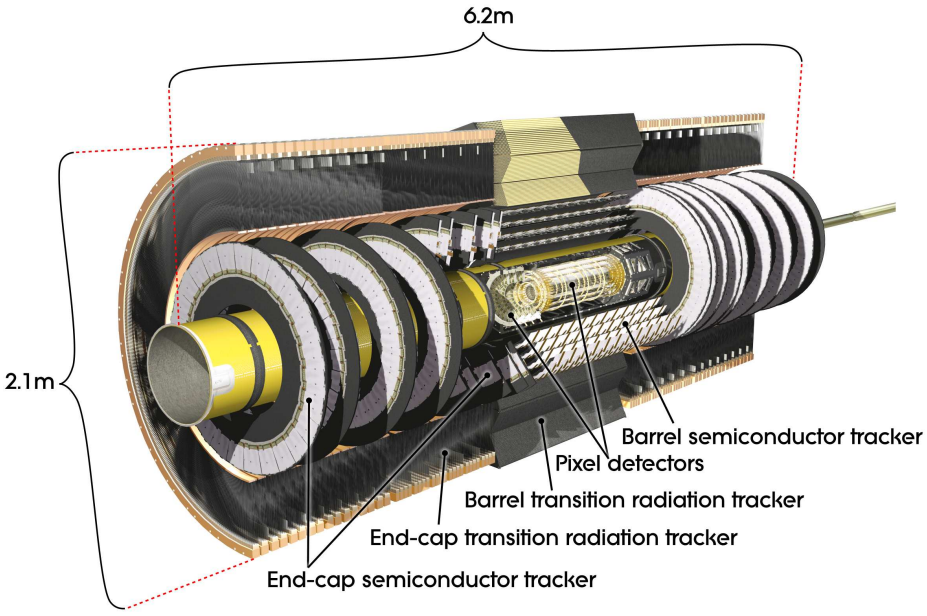


Figure 3.3: Overview of the ATLAS Inner Detector [24].

Semi-conductor Tracker (SCT)

The subsequent ID subsystem is the Semi-conductor Tracker (SCT). It is also a precision silicon detector but the sensors are formed in micro-strips with coarser granularity than pixels but with a sufficiently good spacial resolution (17 μm in $R\phi$ and 580 μm in z or R) in large area at reasonable cost. The micro-strips are arranged in four concentric layers in barrel and in nine disks in each endcap. Each barrel layer is composed of two sub-layers with stereo-strips at 40 mrad angle providing 3D spatial measurements. Total number of readout channels in the SCT is around 6M. This subsystem is an extension of the Pixel system contributing to reconstruction of charged-particle tracks, measurements of particle transverse momenta and reconstruction of vertices.

Transition Radiation Tracker (TRT)

The outermost ID subsystem is the Transition Radiation Tracker (TRT). It is a gaseous tracking detector covering the pseudorapidity range of $|\eta| < 2.0$. It is built of straw tubes, 4 mm in diameter that are parallel to the beam axis in barrel and perpendicular to it in the endcaps. The straws are filled with xenon, carbon dioxide and oxygen mixture. In the TRT, only the $R\phi$ coordinate is measured and the spatial resolution is 130 μm. It is not nearly as precise as the silicon sub-detectors, yet it provides additional 36 (typically) space-points for each traversing charge particle over the radial distance up to over 1 m from the beam axis, greatly improving tracking and momentum measurement resolution. In addition, thanks to its design and the use of xenon, it is capable of providing discrimination between electrons and charged pions with the help of differences in transition radiation induced by those particles.

3.2.3 Calorimeters

The calorimeter system has to be able to measure energies of various types of particles and objects, both charged and neutral in a large range of pseudorapidity. It is therefore built of components providing such capabilities as well as a good spatial resolution. The angular coverage and homogeneity of the calorimeters allow for reconstruction of the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, a quantity often associated with particles escaping detection, such as neutrinos or hypothetical neutralinos.

The overview of ATLAS calorimetry is presented in Fig. 3.4. The calorimeters of the

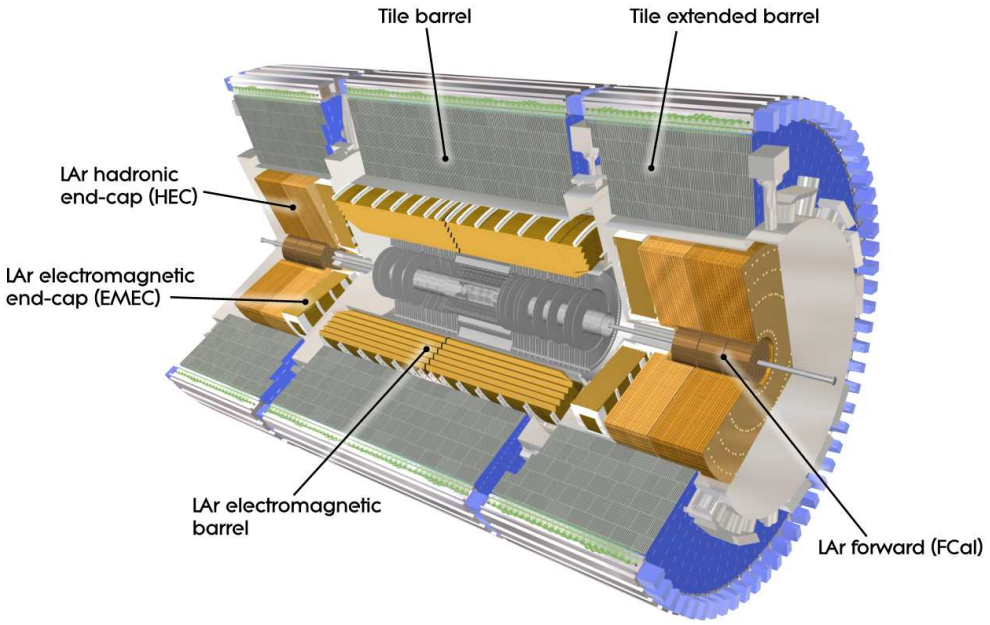


Figure 3.4: Overview of the ATLAS calorimeters [24].

ATLAS experiment are sampling calorimeters, segmented both transversally and longitudinally in so-called cells grouped in layers. The granularity of the segmentation is finest in the layers closer to the interaction point (allowing for studies of electrons and photons) and gets coarser in the outside layers. This allows for studying the shapes of particle showers induced in the calorimeter and improves both the energy calibration and particle identification.

The first (looking from a perspective of a particle going out of a collision) layers build-up the electromagnetic (EM) calorimeter component. It provides pseudorapidity coverage up to $|\eta| < 3.2$ and its fine-granularity region matches the acceptance of the ID ($|\eta| < 2.5$). The EM calorimeter uses liquid argon (LAr) as the active material and lead plates as absorbers to achieve stable and fast response to incoming particles. The finest-granularity first layer has a spatial resolution of 0.025×0.1 in $\eta \times \phi$ which allows for distinction of two photons in a $\pi^0 \to \gamma\gamma$ decay. There is a transition region between barrel and endcaps, at $1.37 < |\eta| < 1.52$ where the performance is worse due to larger amount of detector material. In total there are three layers of the EM calorimeter completed with a thin presampler layer mounted in front of first layer to provide energy sampling to estimate

the loss of energy of electrons and photons in the preceding detector material. The total thickness of the EM calorimeter in ATLAS is 22 to 24 radiation lengths to provide good containment of electromagnetic showers.

The EM calorimeters are followed by Hadronic ones that are suited for measurements of strongly-interacting particles. Their central part consists of plastic scintillating tiles with steel absorber up to $|\eta| < 1.7$. Hadronic endcaps (HEC) and Forward Calorimeters (FCal) use LAr and copper absorbers and cover the pseudorapidity ranges of $1.5 < |\eta| < 3.2$ and $3.1 < |\eta| < 4.9$, respectively. The first layer of FCal is optimised for electromagnetic shower measurements. The total thickness of the ATLAS calorimeter in terms of interaction lengths is around 10 and is sufficient for measurements of high-energy hadronic jets.

3.2.4 Muon Spectrometer

The outermost layer of sub-detectors in ATLAS is the Muon Spectrometer (MS) designed to measure tracks of muons with high momenta (above 3 GeV). It is equipped with a large super-conducting air-core toroid magnet providing magnetic field between 0.5 T and 2 T to enable magnetic deflection of muon tracks and allow for precise transverse momentum measurements. The MS is instrumented with both trigger and precision-tracking chambers. The overview of the MS is shown in Fig. 3.5. The MS is organised in three

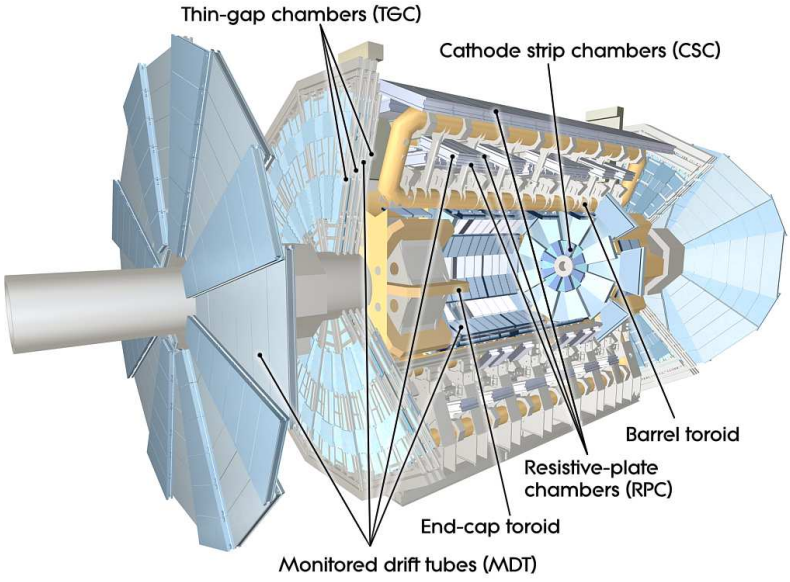


Figure 3.5: Overview of the ATLAS Muon Spectrometer [24].

concentric cylindrical layers in the barrel part and three disks perpendicular to the beam axis in each endcap. The largest part of the tracking coverage is provided by Monitored Drift Tubes (MDT) covering the pseudorapidity range up to $|\eta| < 2.7$ ($|\eta| < 2.0$ in the first layer). These are aluminium tubes filled with argon - carbon dioxide mixture with anode wires in the centre. In the innermost layer at larger pseudorapidities Cathode-Strip Chambers (CSC) are used. They are multi-wire proportional chambers with cathodes segmented into strips. The spatial resolutions of the MDT's and CSC's is 35 μm and 40 μm per chamber respectively. Triggering capabilities are provided for muons within

$|\eta| < 2.4$ by Resistive-Plate Chambers (RPC) in barrel and Thin-Gap Chambers (TGC) in the endcaps. In addition, the trigger chambers also measure the second coordinate of muon position, orthogonal to the one provided by MDT's and CSC's.

3.2.5 Luminosity and Forward detectors

The high-$|\eta|$ region, commonly named the *forward* region in ATLAS is covered by several subsystems. The LUCID (Luminosity measurement using Cerenkov Integrating Detector) is situated at $z = \pm 17$ m (coverage of $5.6 < |\eta| < 6.0$) from the interaction point and detects inelastic pp scattering in the forward region that can be used of online relative luminosity determination.

The ALFA (Absolute Luminosity For ATLAS) is located at $z = \pm 240$ m from the interaction point and consists of Roman-pot detectors built-up with scintillating fibres. The Roman pots are approaching beam at 1 mm distance in case operation conditions (special dedicated beam optics, reduced beam emittance) are fulfilled. ALFA measures absolute luminosity by means of elastic pp scattering. Its amplitude in the forward direction can be used to extract the absolute luminosity.

The Zero-Degree Calorimeter (ZDC) plays a key role in determining the centrality of heavy ion collisions in ATLAS. It is located at $z = \pm 140$ m where the common vacuum beam pipe splits into two separate ones. The ZDC is built of tungsten layers and quartz rods that measure neutral particles at $|\eta| > 8.2$.

The ATLAS Forward Proton (AFP) system [101] consists of two pairs of two detector stations located at $z = \pm 205$ and $z = \pm 217$ m and has been in operation from 2016. It aims for measuring protons scattered at very small angles at the interaction point and does not require any special running conditions, so can collect data together with the rest of the ATLAS detector and provide triggering.

3.2.6 Trigger

The rate of bunch crossings at the LHC is around 40 MHz (with 25 ns bunch-spacing time). It is not possible to store all collision events with average size of 1.3 MB, therefore a reduction of rate down to 200 Hz in Run 1 and 1 kHz in Run 2 had to be made. A fast and efficient triggering is therefore an absolute necessity for the whole operation of ATLAS.

The design of the trigger system is based on multiple triggering levels. The Level-1 (L1) trigger is entirely hardware-based and uses dedicated electronic systems to collect the available information about an event and to perform the decision whether to reject the event or to process it further. The L1 collects event information from only a subset of detectors with fast-enough readout, such as muon trigger chambers (TGC's and RPC's) and reduced-granularity calorimeter towers. These data are then passed on to Central Trigger Processor (CTP) that combines them and computes trigger decision based on the implemented selection criteria known as *trigger menus*. The L1 also constructs Regions of Interest (RoI), the $\eta - \phi$ areas of the detector where L1 has identified interesting features. If the event is accepted by L1, it is passed on to higher trigger levels and the information about RoI's is subsequently used there. The L1 output event rate is reduced to 75 kHz (100 kHz in Run 2 [102]) and the average decision time is 2.5 μs.

The selection at Level-2 (L2) trigger bases on the RoI information provided by L1, however L2 is enabled to use the full detector information from the RoI areas (around 2% of total information on the event). The algorithms are software-based and are running

on a dedicated computer farm. The average decision time per event is 40 ms and the L2 output rate is 3.5 kHz. Events selected by L2 are then further processed at the final Event Filter (EF) level. It is also software-based and has nearly-full offline selection procedures implemented and takes decision in 4 s on average. In Run 2 the L2 and EF have been merged together (and are commonly denoted High-Level Trigger, HLT) to allow for a better sharing of computing resources and simplification of both the hardware structure and software algorithms [102].

3.3 Simulation of collision events

Theoretical understanding of the measured or searched processes and their backgrounds is essential in experimental analyses. Therefore, a common language between the theory and experiment was established by means of the Monte Carlo (MC) simulations. Computer-based generation of pseudo-random numbers convoluted with the mathematical description of the known or hypothetical process of interest are well-suited for modelling or simulating pseudo-events representing such process. Monte Carlo simulations are usually performed in steps starting from generation of physics events at parton level that are further developed to stable-particle level and passed on to detector simulation software. In the ATLAS experiment this is all performed within the Athena software framework [103, 104].

The first stage of production of simulated events is performed by so-called event generators, programs containing theoretical models of the generated processes. A number of such programs exist, including general ones like Pythia [105], Sherpa [106], Herwig [107, 108] and many others. They generate hard interactions of partons, initial and final-state radiation and underlying event. The products of the hard process are then hadronised (if applicable) and prompt decays are performed (this, in principle, can be made with a different generator). The choice of MC generator and, therefore, the theoretical model of a given process, depends on the analysis and on the process under consideration.

A generated event with final-state particles is then propagated through the detector simulation to asses the response of detector subsystems to each particle as well as to simulate the interaction of particles with the material of the detector and its magnetic field. This is obtained with the help of GEANT4 framework [109] using detailed information on the detector geometry, material and magnetic fields. The output of the simulation step is returned in the way the information from real collision events is read-out from detector electronics. The step of translating the simulated events into the language of electric currents and voltages is called *digitization*.

The effects of pile-up interactions are present in ATLAS data (actually, reducing their impact can be challenging), so a proper simulation of this effect is also needed. This is achieved by overlaying simulated Minimum-Bias events on top of hard-interaction ones in an additional step in event production.

The simulated events can then be passed on to the event-reconstruction software infrastructure and can be further processed in exactly the same way as real collision data.

3.4 Reconstruction and identification of physics objects

The process of transforming the raw detector output, both from simulated and real collision events into meaningful information about particles and physics objects (such as

jets) is called *reconstruction*. In this process all basic-level detector information like hits in tracking detectors or cells in calorimeters are collected and dedicated software algorithms convert them first into more complex objects like tracks or calorimeter clusters and then, in the second step into particle and object candidates. The candidates are then subject to *identification* that, based on the candidate specifics and the available information, classifies the given candidate as a particle or object of a given type. This classification, in principle, can be more or less restrictive, depending on the needs of particular physics analysis. A less restrictive classification is more efficient for genuine particles of given type but more often mis-classifies other objects as objects of investigated types. A more restrictive identification results in lower rate of mis-classified objects for a price of lower classification efficiency for real objects of given type. An overview of reconstruction and identification procedures of physics objects in the ATLAS experiment is given below.

The reconstruction of electrons [110–112] is seeded by the information from the EM calorimeter. Cells with significant energy deposits are grouped together in clusters if they are not spatially separated. The clusters are then matched to tracks extrapolated from the Inner Detector to suppress photons. This procedure is highly efficient reaching 95% efficiency already for electrons with transverse energy $E_{\mathrm{T}} > 7$ GeV and 99% at $E_{\mathrm{T}} > 45$ GeV. Not all the reconstructed objects are real electrons however. A large fraction of QCD jets is also reconstructed as electron candidates. Moreover, rejection of electrons from photon conversions and Dalitz decays is also needed. This is handled by the identification step. It takes into account the longitudinal and lateral shapes of EM calorimetric clusters, the properties of tracks in the ID (i.e. number of hits in Pixel and SCT detectors, transition radiation in the TRT) and the goodness of matching between tracks and clusters. Both simple cut-based and advanced multivariate discriminants are used, with three working points, *loose*, *medium* and *tight*, defined for each of them. The working points are defined in such way that the electron candidates fulfilling *tight* selection criteria are a subset of electron candidates fulfilling *medium* criteria which, in turn, are a subset of *loose* electron candidates. This convention is commonly used in the whole ATLAS experiment. Additionally, a dedicated algorithm reconstructs and identifies forward electron candidates outside of the ID acceptance in the $2.5 < |\eta| < 4.9$ region using calorimetric information only.

Photons [113] are reconstructed from the EM calorimeter energy deposits similarly to and in parallel to electrons with the exception of track-cluster matching procedure. An unconverted photon has no track matched to calorimetric cluster whereas a converted one can have one or two tracks, consistent with conversion vertex within the ID volume, matched to the cluster. Photon identification bases on the cluster-shape variables and energy leakage into hadronic calorimeter.

The reconstruction of muons [114, 115] uses the information from the Muon Spectrometer and ID. Muons with low transverse momenta, $p_{\mathrm{T}} < 3$ GeV do not reach the MS because of the amount of detector material between the interaction point and the MS. Such muons are very hard to reconstruct as they do not produce any significant signal in the spectrometer. Muons with $p_{\mathrm{T}} > 3$ GeV can be reconstructed in several ways but a most commonly-used reconstruction algorithm combines the tracks reconstructed in the ID and MS independently, resulting in highest-purity samples. In Run 2 an identification algorithm was applied to reduce the background from pion and kaon decays and to select prompt muons. This algorithm applies quality requirements on the combined ID-MS tracks.

The QCD-coloured objects, quarks and gluons, frequently produced in the LHC col-

lisions cannot exist as free particles after production. They therefore undergo parton showering (producing additional coloured objects) and hadronisation followed by hadron showering before reaching the calorimeters. These showers of particles usually appear in the calorimeters as collimated energy deposits. Such objects are called *jets* and are among the most important physics objects for analyses at the LHC. Ideally, a single jet could approximate the four-momentum and flavour of the quark or gluon of origin. A number of algorithms for constructing jets from calorimeter clusters exist, including the anti-k_{T} [116], most commonly used in the ATLAS experiment with a certain distance parameter R[2], usually 0.4 or 1.0. The algorithm uses so-called topological clusters [117] as inputs and in each event a set of numbers is defined: $d_i = 1/p_{\mathrm{T},i}^2$ and $d_{i,j} = \min(1/p_{\mathrm{T},i}^2, 1/p_{\mathrm{T},j}^2) \times \Delta R/R$ (with i, j running over topological clusters). If the minimum of this set is found to be one of the d_i's, this cluster is removed from further processing and classified as a separate jet. If, in contrary, it is found to be one of the $d_{i,j}$'s, the i-th and j-th clusters are merged and added back to the list of algorithm inputs. The iterations continue until no more input objects are present. In this way, the energy deposits are combined around highest-energetic cluster but with the distance between neighbouring jets kept at least at the level of the R parameter. This algorithm is both infrared- and soft-emission-safe. Reconstructed jets are subject to special energy-scale calibration procedures that are based on MC simulations and correct jet energies to better match the true values, accounting for detector features, jet-reconstruction algorithms and jet fragmentation. The default calibration assumes electromagnetic nature of calorimeter clusters (so-called EM-scale calibration) but dedicated corrections, such as Local Cell Weighting (LCW) [118], can be applied to clusters classified as hadrons to account for non-compensating nature of the ATLAS calorimeters. In addition, pile-up and jet-origin corrections are applied at this stage [119].

In the proton – proton collisions at the LHC the beams collide nearly head-on and therefore the initial state has no transverse momentum. Therefore the non-zero vector sum of transverse momenta of final-state objects may indicate the presence of undetected particles produced in a collision. They can not only be the SM neutrinos but also hypothetical particles from BSM scenarios, such as neutralinos or Dark Matter particles. Because of that, the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$ (often referred to as missing transverse energy) is an important experimental observable for both SM measurements and New Physics searches. Its reconstruction involves all detector subsystems and is limited by acceptance and signal from pile-up. The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction in ATLAS [120] bases on two classes of contributions. First is the so-called hard component related to fully-reconstructed jets and particles (electrons, muons, photons and τ-leptons). As the hard objects are reconstructed independently, an overlap between them occurs frequently. Dedicated ambiguity-resolving procedures are therefore in place, usually with a preference for electrons over photons and over τ leptons with jets at the last place (muons usually have very little overlap with other objects being reconstructed from ID and MS). The second component, called soft signal, results from the existence of charged-particle tracks that are associated with the primary (hard-process) vertex but not with any of the hard objects.

The reconstruction and identification of the τ leptons as well as flavour-tagging of jets initiated by heavy quarks is described in detail in Chapter 4.

[2]The R parameter is defined in the same way as the ΔR distance and describes the typical maximal radius of a jet in $\eta - \phi$ coordinates.

3.5 Luminosity measurements

The delivered luminosity needs to be precisely known for many of the physics analyses performed in the experiment. The uncertainty on its measurement translates directly into to the uncertainty on cross-section measurements. Also in the searches for new phenomena, the accurate determination of the integrated luminosity is needed to evaluate the levels of backgrounds and assess the search sensitivities.

The ATLAS detector monitors the luminosity by measuring the rate of interactions per bunch crossing using the LUCID and BCM (Beam Conditions Monitor) detectors and by monitoring the multiplicities of charged particles produced in randomly-selected bunch crossings [96]. This is then subject to the van-der-Meer calibration [121] using beam-separation scans. The resulting uncertainties of luminosity determination in Run 1 are $\pm 3.5\%$, $\pm 1.8\%$ and $\pm 1.9\%$ for the 2010, 2011 and 2012, respectively. Similar methodology was in place for Run-2 data taking and the resulting uncertainty of the determination of integrated luminosity is $\pm 2.1\%$ for the combined 2015 and 2016 periods. In the combined (2015-2018) Run-2 dataset the luminosity uncertainty is $\pm 1.7\%$.

3.6 Quality of data for physics analyses

The overall quality of the collected data is assessed nearly on-the-fly during the data-taking periods. For most of the physics analyses performed in the ATLAS collaboration a common set of quality criteria exists based on the operation of all subsystems. Usually only the events with all subsystems fully operational are accepted for further studies. Lists of such periods, so-called runs (of the order of hours) and luminosity blocks (of the order of minutes) are created and events from outside of them are rejected. Additionally, event-level cleaning is also performed to reduce the impact from electric discharges in the hadronic calorimeters, coherent noise, cosmic rays or beam-background. They could, in principle, mimic high-energy objects in calorimeters or high-energy tracks in MS or ID and artificially create jets or charged particles and distort the measurements of $E_{\mathrm{T}}^{\mathrm{miss}}$. Also, non-functioning parts of sub-detectors are accounted for in the reconstruction to recover objects and particles that would not be able to fulfil their reconstruction criteria otherwise.

Chapter 4

Reconstruction of hadronic final states associated to the decays of τ leptons and b quarks

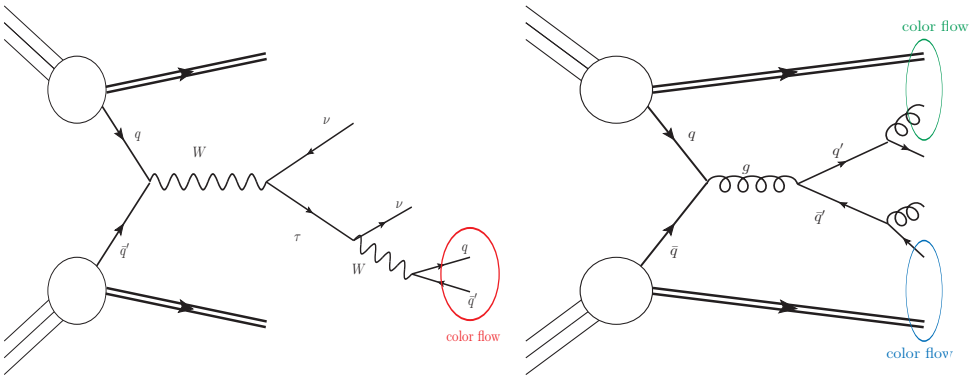
The most abundant class of processes occurring in proton – proton collisions at the LHC is the production of QCD multi-jets. It is therefore a very challenging task to attempt to reconstruct and identify complex objects that manifest their presence in the experiment by producing special types of jets. Hadronically-decaying τ leptons produce, in principle, narrow and collimated jets with relatively low multiplicities and a significant electromagnetic component due to the presence of π^0's. Nevertheless, the τ production rates are overwhelmed by those of multi-jets by many orders of magnitude. Jets resulting from hadronisation and decay of b-quarks on the other hand have usually displaced production vertices due to a significant b-hadron lifetime. Serving as probes for the top and Higgs sector (both SM and BSM) they are also subject to huge contamination from light-jet background.

This chapter describes the reconstruction and identification of τ leptons in the ATLAS experiment as well as techniques used for tagging of heavy-flavour jets.

4.1 Reconstruction and identification of hadronic decays of τ leptons

The τ lepton has a relatively short lifetime (see Section 2.1) and decays before reaching any of the ATLAS sub-detectors. Therefore, despite still being an elementary lepton, it cannot be observed as a simple particle (such as electron or photon) but has to be reconstructed as a complex object instead. In 35% of cases it decays into a lighter charged lepton and two neutrinos. This class of decays is very difficult to distinguish from prompt lepton production so the actual τ reconstruction considers only hadronic decays of τ leptons occurring in 65% of the cases. Additionally, no attempt is made to reconstruct ν_τ in such decays, only the so-called visible part of the hadronic decay can be reconstructed.

The properties of τ-initiated jets are significantly different from those of quark- and gluon-initiated ones. This is because of the difference in colour structure of these jets. The τ's are colour-less and decay into a neutrino and a virtual W-boson that further decays into a pair of colour-connected quarks. The QCD jets are initiated by a coloured object - quark or gluon. Therefore, a colour connection between the jet and beam remnants can exist. Such colour connection has sufficient energy to produce additional quark - anti-quark pairs in a broader cone around the original coloured parton (closer to beam remnants)[1], see Fig. 4.1.



(a) (b)

Figure 4.1: Schematic view of colour flow in hadronic decays of τ leptons (left) and in QCD jets (right). The former shows $W \to \tau\nu$ production mode as an example but can be generalised to other modes (in $Z \to \tau\tau$ the first neutrino can be exchanged for a τ lepton, in $t\bar{t}$ decays the picture is more complex but conclusions regarding τ-jets remain unchanged). The QCD example shows gluon-initiated jet, quark-initiated jet develops analogously and can be imagined by exchanging quark lines for gluon lines and vice-versa.

The reconstruction and identification of hadronic τ decays, even though based on similar principles and features, was evolving during past years of LHC operation. In the following both the Run-1 baseline τ reconstruction [123] and the improvements for

[1] A concept of dragging a gluon-like string between the beam remnant and parton in jet can serve to help imagining such behaviour. With sufficient energy the string breaks creating a quark - anti-quark pair outside of the jet centre, between the considered parton and beam-remnant. This concept was validated by experimental observations of hadronically decaying W bosons from $t\bar{t}$ decays at D0 experiment [122].

Run 2 [124, 125] are discussed.

4.1.1 Features of τ reconstruction in ATLAS

The τ-reconstruction algorithm is seeded with jets, formed by anti-k_{T} algorithm with the distance parameter $R = 0.4$, using topological clusters as an input. Seed jets are required to have their transverse momenta $p_{\mathrm{T}} > 10$ GeV and to be contained within the ID acceptance, $|\eta| < 2.5$. To increase the reconstruction efficiency and reduce the effects of pile-up a tau-vertex finding is then performed with the help of tracks with $p_{\mathrm{T}} > 1$ GeV located in the core region of the seed jet, $\Delta R < 0.2$ from the jet direction. A vertex with the largest matched fraction of the sum of track transverse momenta is chosen as the τ production vertex and is further used to reconstruct τ direction, associate tracks and calculate identification variables.

The momentum of the τ candidate is obtained from four-momenta of topological clusters constituting the seed jet after recalculating the η and ϕ of their barycentre with respect to the τ production vertex found in the previous step. The mass of the reconstructed candidate is defined to be zero.

Afterwards, track association with τ candidate is performed. Tracks are assigned to the candidate if they are located in the core $\Delta R < 0.2$ region around the candidate flight direction. Their minimum transverse momentum is 1 GeV and required number of hits in the silicon trackers is at least 7. Moreover, it is required that the transverse ($|d_0|$) and longitudinal ($|z_0 \sin\theta|$)[2] impact parameters with respect to τ vertex are no larger than 1 mm and 1.5 mm, respectively. Tracks selected in this manner are then used for classification of each τ candidate based on its track multiplicity. Two classes of candidates exist, single-prong and three-prong, with one or three charged tracks, respectively. Single-prong and three-prong decays constitute about 77% and 23% of hadronic τ decays, respectively.

In addition to the above, an algorithm for finding π^0 mesons in τ-jets is applied to each τ candidate. The algorithm, Pi0Finder, consists of two steps. First, a determination of the number (zero, one or two-or-more) of π^0's is made with the help of global candidate characteristics. Boosted Decision Trees (BDT) [126, 127], machine-learning algorithms, are employed to provide such classification. They use various combinations of track and cluster information as an input, namely:

- Fraction of $\pi^{\pm}$ energy in the EM calorimeter obtained as a fraction of the difference between sum of track momenta and energy in hadronic calorimeter, divided by total τ-candidate energy in the EM calorimeter,
- τ-lepton energy deposed in the EM calorimeter divided by total energy of all associated tracks (different from the above as no track quantities are used in the numerator),
- Fraction of energy in the pre-sampler layer of calorimeter with respect to the total τ energy,
- Number of cells associated to τ-clusters with energy above 200 MeV,
- Ratio of τ transverse energy to the transverse momentum of the leading track.

[2]The $\sin\theta$ factor is introduced to reduce the track-selection efficiency dependence on the track polar angle.

The counting is then performed by two BDTs with the same set of input variables. One is responsible for deciding whether the τ-candidate decay contains any π^0's at all and the other one distinguishes the 1 π^0 and 2+ π^0's cases and is only used in single-prong decays. The efficiency is much higher for the former, reaching almost 90% of correctly classified 1p decays with no π^0's and 80% correctly-classified 1p decays with at least 1 π^0 present. In the 3p case, the efficiency of 3p-0π^0 classification exceeds 93%, although the 3p events with π^0's are very often misclassified as containing no π^0's (over 50% of the cases). On the other hand, similar misclassification efficiency is present in the 1p decays where it comes to distinguishing 1π^0 from $\geq$ 2π^0 decays. Detailed efficiency matrix of the π^0-counting cluster-based algorithm is shown in Fig. 4.2 [124]. The efficiency for correctly classifying a given decay mode with this algorithm is 68.9%.

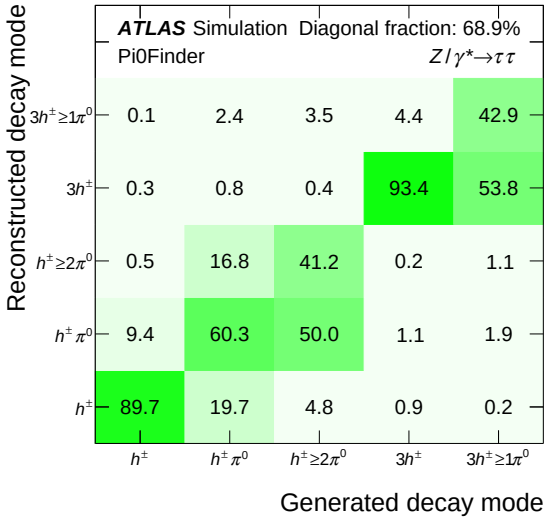


Figure 4.2: Efficiency matrix for decay-mode classification by the cluster-based π^0-counting algorithm in simulated $Z \to \tau\tau$ events, showing the probability that a given generated τ decay mode will be classified as particular mode. No decays with neutral kaons are considered. The τ leptons have to fulfil basic selection criteria: $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. The statistical uncertainty is negligible [124].

In the second step, independently from π^0 counting results, reconstruction of π^0 kinematics is performed. Pile-up and noise activity is estimated in the ring of $0.2 < \Delta R < 0.4$ around the τ direction and this activity is then subtracted from clusters within the $\Delta R < 0.2$ core cone. Then, a cluster or a pair of clusters is selected based on a so-called π^0-likeness score S_{π^0}, defined in Eq. 4.1.

$$S_{\pi^0} = \frac{E^0_{\mathrm{cluster(s)}}}{f^{\mathrm{HAD}}_{\mathrm{cluster(s)}} + x\sqrt{\frac{E_{\mathrm{cluster(s)}}}{E_{\mathrm{calo}} - p_{\mathrm{tracks}}}}} \tag{4.1}$$

In the above, $E^0_{\mathrm{cluster(s)}}$ is the energy of the cluster(s) in the pre-sampler layer of calorimeter, $f^{\mathrm{HAD}}_{\mathrm{cluster(s)}}$ is the fraction of cluster(s) energy in the hadronic calorimeter, $\frac{E_{\mathrm{cluster(s)}}}{E_{\mathrm{calo}} - p_{\mathrm{tracks}}}$ is the ratio of cluster(s) energy to the simple measure of π^0 energy (difference between total

τ energy and track momentum representing the total momentum of charged pions). The x parameter is chosen based on optimization of the score-based selection on simulated τ decays and its value is set to 0.4. Afterwards, a correction to cluster energy is applied to reduce the effects of double-counting of π^0 energy. It is especially important at higher p_{T} of τ leptons due to merging of calorimeter clusters and higher hadronic contamination in the reconstructed π^0 clusters.

The properties of the reconstructed π^0 clusters together with the information on τ tracks end energy deposits is then used in the identification of τ candidates, described in the next Subsection.

The approximate overall percentage of the reconstructed hadronic τ decays including detector and kinematic acceptance (efficiency times acceptance) is between 32% and 48% (depending on the number of neutral pions) for single-prong decays and around 38% for three-prong decays [124].

4.1.2 Identification of hadronically-decaying τ leptons in Run 1

Despite the fact that the τ-lepton reconstruction procedures base on the properties of hadronic τ decays, they provide very little rejection against QCD background jets. It is the identification step that is responsible for the actual suppression of jets. This step uses detailed information from calorimeter and tracker both in the core and in the isolation regions of the reconstructed τ candidate, including the reconstructed π^0 content. The variables used in Run-1 τ lepton identification in ATLAS are listed below [123]. The relevant Figures present their distributions for both the simulated signal (τ leptons from the decays of W, Z, simulated with Alpgen [128] interfaced with Pythia 6 [129], and hypothetical Z' bosons simulated with Pyhtia 8 [105]) and background obtained from ATLAS Run-1 data.

- Central energy fraction (f_{cent}): Fraction of transverse energy in the central $\Delta R < 0.1$ cone around the τ candidate with respect to the total energy in the core cone ($\Delta R < 0.2$) of the τ candidate. The energy is obtained from all cells belonging to topological clusters with barycentre in this region, calibrated at EM scale. Pile-up corrections based on the number of primary vertices in the event are applied to reduce pile-up-related biases. This variable is used in both single-prong (see Fig. 4.3 left) and three-prong (Fig. 4.7 left) τ identification.

- Leading track momentum fraction (f_{track}): The transverse momentum of the charged particle with highest p_{T} in the core region of the τ candidate, divided by the sum of transverse energy in all cells belonging to topological clusters in that region, calibrated at EM scale. A correction depending on the number of primary vertices is applied to this variable to make it pile-up independent. The f_{track} is used in the identification of both 1-prong (Fig. 4.3 right) and 3-prong (Fig. 4.7 right) candidates.

- Track radius (R_{track}): Sum of p_{T}-weighted distance of all τ-associated tracks to the τ candidate direction. It is used in both the single-prong (Fig. 4.4 left) and three-prong (Fig. 4.8 left) cases.

- Mass of the track+π^0 system ($m_{\pi^0+\mathrm{track}}$): Invariant mass of the system composed of tracks and π^0 mesons in the core region, used both for 1-prong (Fig. 4.4 right) and 3-prong (Fig. 4.8 right).

- Number of π^0 mesons (N_{π^0}): Number of reconstructed π^0 mesons in the core region. This information is used for both 1-prong (Fig. 4.5 left) and 3-prong (Fig. 4.9 left) identification.

- Ratio of track+π^0-system p_{T} ($p_{\mathrm{T}}^{\pi^0+\mathrm{track}}/p_{\mathrm{T}}$): Ratio of transverse momentum calculated from tracks and reconstructed π^0 mesons to the p_{T} of the τ candidate. This variable is the last one used in identification of both track-multiplicity cases of τ candidates (single-prong: see Fig. 4.5 right, three-prong: Fig. 4.9 right).

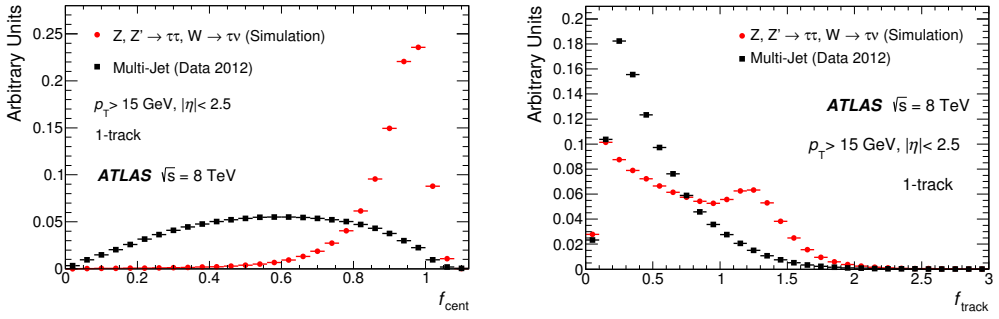
- Leading track impact-parameter (IP) significance ($S_{\mathrm{leadtrack}}$): Transverse impact parameter of the highest-p_{T} track in the τ core region (with respect to τ vertex) divided by its estimated uncertainty. It is used in the identification of single-prong τ candidates only (Fig. 4.6 left).

- Number of tracks in the isolation region ($N_{\mathrm{track}}^{\mathrm{iso}}$): Number of tracks associated with the τ candidate in the isolation ring $0.2 < \Delta R < 0.4$ around τ direction. Only used in single-prong τ identification (Fig. 4.6 right).

- Maximum ΔR (ΔR_{Max}): Maximum cone distance between a core-region track associated to the τ candidate and its direction. This variable is used in 3-prong τ identification only (Fig. 4.10 left).

- Transverse flight path significance ($S_{\mathrm{T}}^{\mathrm{flight}}$): Decay length of secondary vertex, reconstructed from τ-associated tracks in the core region, in the transverse plane with respect to the τ vertex, divided by its estimated uncertainty. It is defined for 3-prong candidates only (Fig. 4.10 right).

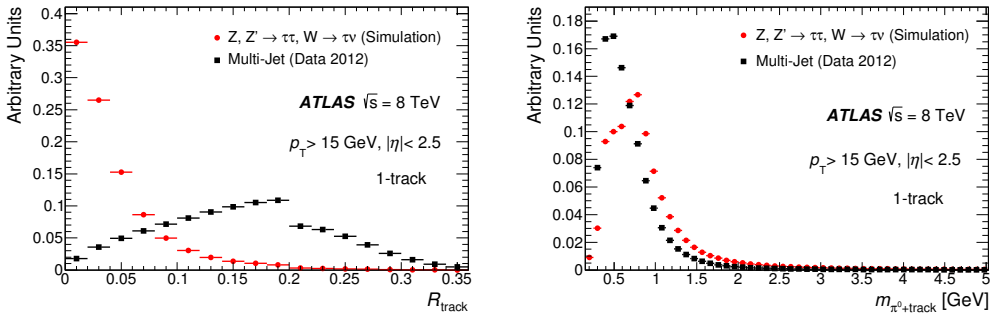
- Track mass (m_{track}): Invariant mass calculated from four-vectors of all core and isolation-region tracks associated to the τ candidate, assuming pion mass for each track. It is used in the identification of three-prong τ candidates only (Fig. 4.11).



(a) (b)

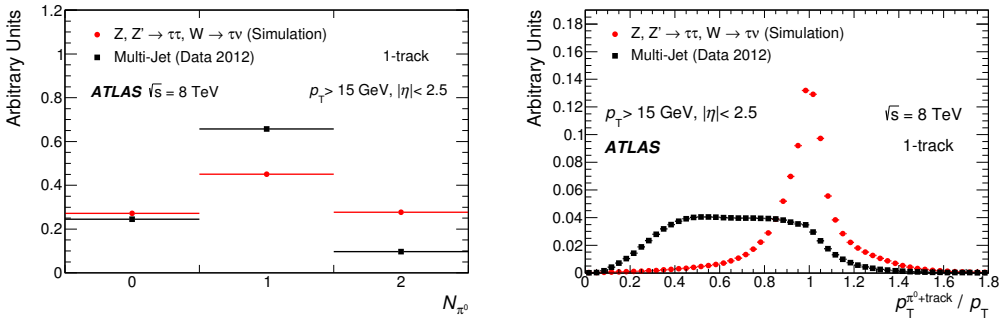
Figure 4.3: Signal and background distributions of the single-prong τ identification variables: f_{cent} (left) and f_{track} (right) [123].

The variables listed above are used to construct two separate BDT algorithms, one for each track-multiplicity category. These algorithms are trained with simulated τ leptons in



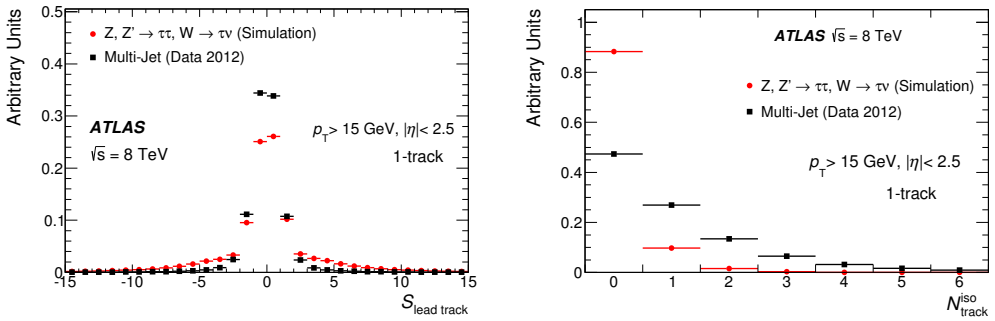
(a) (b)

Figure 4.4: Signal and background distributions of the single-prong τ identification variables: R_{track} (left) and $m_{\pi^0+\text{track}}$ (right) [123].



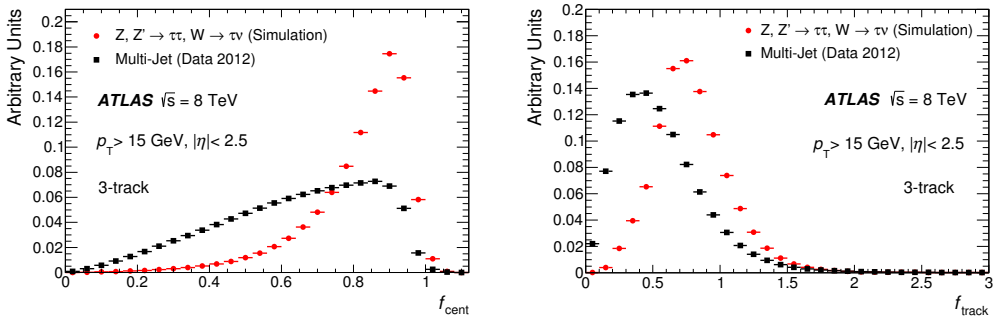
(a) (b)

Figure 4.5: Signal and background distributions of the single-prong τ identification variables: N_{π^0} (left) and $p_{\text{T}}^{\pi^0+\text{track}}/p_{\text{T}}$ (right) [123].



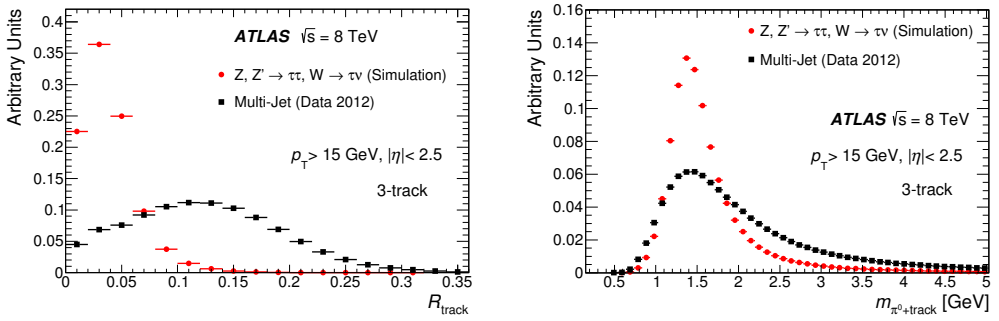
(a) (b)

Figure 4.6: Signal and background distributions of the single-prong τ identification variables: $S_{\text{leadtrack}}$ (left) and $N_{\text{track}}^{\text{iso}}$ (right) [123].



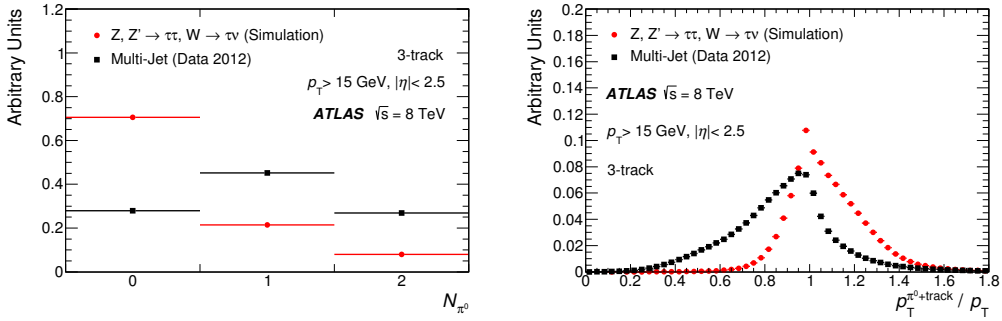
(a) (b)

Figure 4.7: Signal and background distributions of the three-prong τ identification variables: f_{cent} (left) and f_{track} (right) [123].



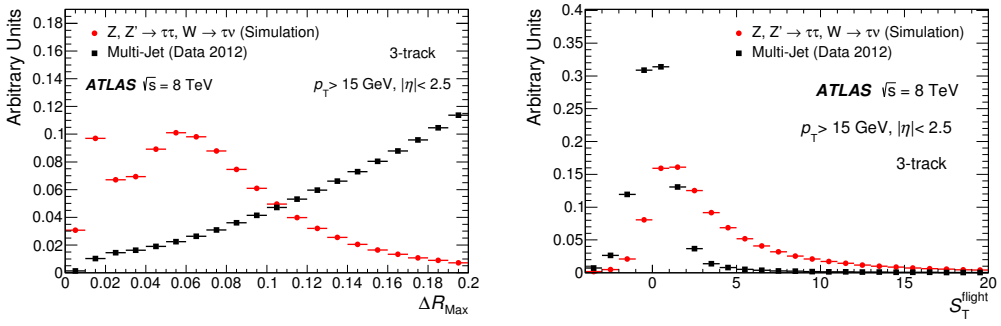
(a) (b)

Figure 4.8: Signal and background distributions of the three-prong τ identification variables: R_{track} (left) and $m_{\pi^0+\text{track}}$ (right) [123].



(a) (b)

Figure 4.9: Signal and background distributions of the three-prong τ identification variables: N_{π^0} (left) and $p_{\text{T}}^{\pi^0+\text{track}}/p_{\text{T}}$ (right) [123].



(a) (b)

Figure 4.10: Signal and background distributions of the three-prong τ identification variables: ΔR_{Max} (left) and $S_{\text{T}}^{\text{flight}}$ (right) [123].

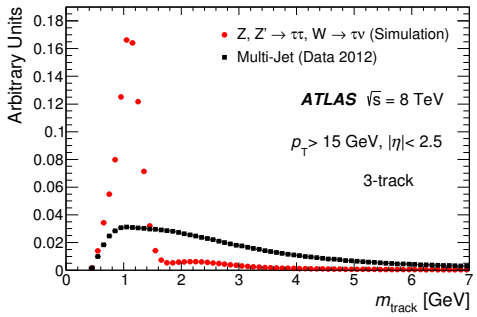
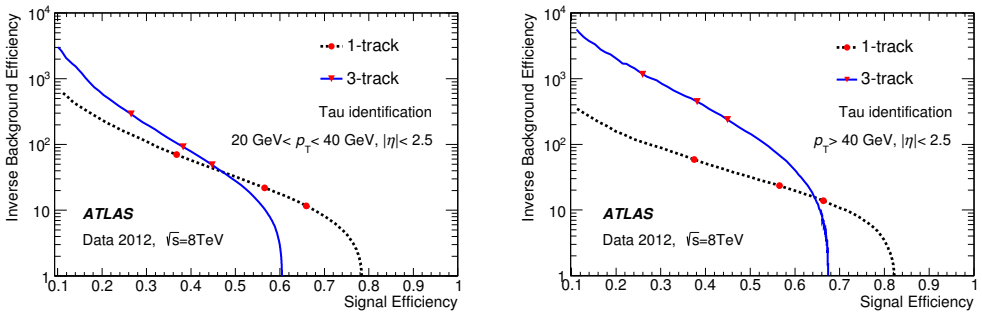


Figure 4.11: Signal and background distributions of the three-prong τ identification variable: m_{track} [123].

W, Z and Z' decays. Background events are obtained from real Run-1 ATLAS data selected with jet triggers. The fraction of real τ leptons in these events is negligible. In the signal sample, the reconstructed hadronically-decaying τ leptons are required to be matched within $\Delta R < 0.2$ cone to generator-level (*true*) τ's with transverse momentum of the visible part of the τ (i.e. excluding neutrino) of $p_{\text{T,vis}}^{\text{true}} > 10$ GeV and pseudorapidity of $|\eta_{\text{vis}}^{\text{true}}| < 2.3$. The efficiency for signal used in the following part can be understood as the fraction of true hadronic decays of τ's with a given number of charged tracks which are reconstructed as candidates with that number of tracks and which fulfil the identification criteria. The background efficiency on the other hand is defined as the fraction of reconstructed and identified τ decays with a given number of charged tracks in the background sample.

As can be seen from Figs. 4.3 – 4.11, the identification variables do not provide perfect signal-background separation and there is always an overlap between them. In other words, there is effectively no non-zero signal-identification efficiency point with 100% rejection (zero efficiency) for background. Therefore, specific signal-efficiency points with known background contamination are chosen as so-called working points for analyses with τ leptons. The requirements on the BDT score (output from the identification algorithm) are chosen in such way that the resulting efficiency is stable with respect to true τ-p_{T}.

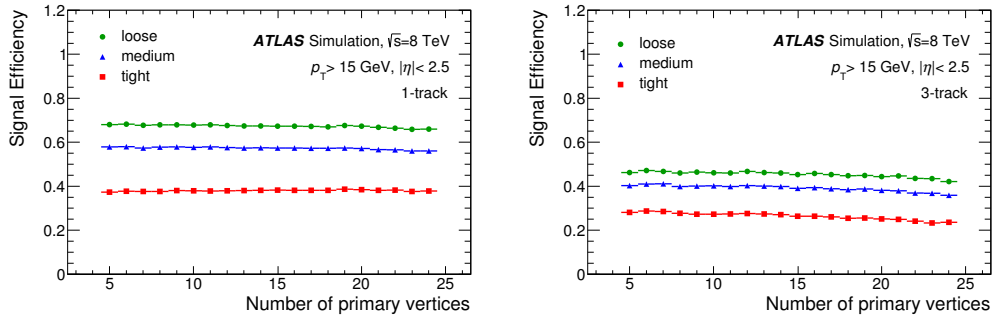
Here, similarly to the case of electrons, the working points are called *loose*, *medium* and *tight*. The *tight* selection criteria result in lower signal selection efficiency with a low contamination from background events in the selected sample. Inversely, *loose* selection results in more accepted signal events and higher background contamination. It is up to the needs of a specific analysis which working point to choose. Fig. 4.12 presents the performance of the identification algorithm in terms of inverse background efficiency versus signal efficiency in two τ-p_{T} regimes. Thanks to the choice of identification variables (and



(a) (b)

Figure 4.12: Inverse background efficiency versus signal efficiency for low-p_{T} (left) and high-p_{T} τ candidates (right) in Run 1. The red markers correspond to the three working points (see text) [123].

applied corrections), the performance of the identification algorithm does not degrade with increasing number of pp collisions in bunch-crossings, as shown in Fig. 4.13.



(a) (b)

Figure 4.13: Signal efficiency for the three identification working points (see text) for single-prong (left) and three-prong (right) τ candidates with respect to the number of primary vertices in the event [123].

4.1.3 Suppression of electrons and muons in Run 1

In addition to the most apparent background objects - the QCD jets, the hadronic decays of τ leptons, especially of the 1-prong type, can also be mimicked by electrons and muons. Dedicated algorithms have therefore been developed to suppress τ candidates reconstructed from these objects.

Electrons constitute the sub-leading class of background object for single-prong τ candidates. However, despite many similarities, there are some characteristics that help distinguishing these two types of objects. The *electron veto* algorithm is, again, a BDT that uses a set of input variables known to discriminate between τ's and electrons, namely:

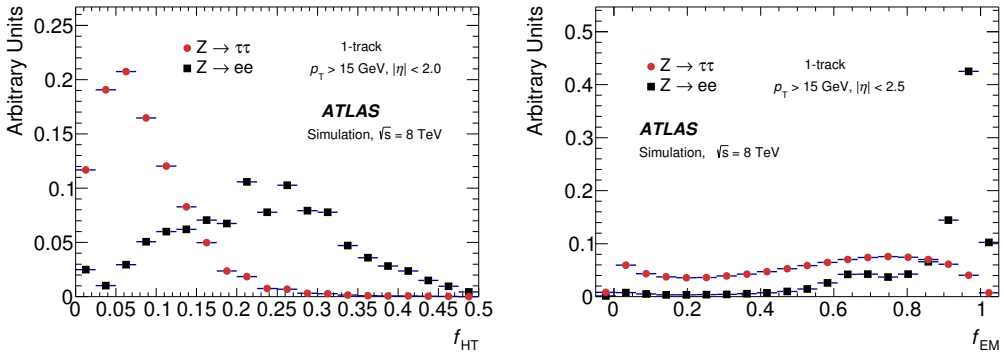
- Ratio of high-threshold to low-threshold hits in the TRT, f_{HT}, describing the fact that electrons are more-likely to induce transition radiation than the pions are,
- Angular distance between the track and the direction of τ candidate from the calorimeter,
- Ratio of energy deposited in the EM calorimeter to the energy deposited in the hadronic calorimeter, f_{EM}, much larger for electrons,
- Amount of energy leaking to hadronic calorimeter,
- Ratio of energy deposited in the $0.1 < \Delta R < 0.2$ to the energy in the $\Delta R < 0.2$ cone around τ candidate direction.

The distributions of f_{HT} and f_{EM}, two best-separating variables, are presented in Fig. 4.14. The performance of electron-veto algorithm is presented in Fig. 4.15. In both these Figures the electron background is obtained from the simulation of $Z \to ee$ process and signal consists of $Z \to \tau\tau$ events with reconstructed candidates matched to generator-level τ's as previously. Slight differences of electron-veto performance are observed with respect to η region due to the fact that the f_{HT} variable is only available for $|\eta| < 2.0$ and that separate optimisation was performed in different pseudorapidity regions. Again, three working points are defined for the electron veto, *loose*, *medium* and *tight* with signal efficiencies of 95%, 85% and 75%, respectively.

Most of the hadronic τ decay candidates mis-reconstructed from muons can be suppressed by standard muon-reconstruction algorithms. Reduction of the remaining contribution is performed by a simple cut-based *muon veto* algorithm. Muons that deposit large amounts of energy in the calorimeter fail standard muon reconstruction. They are characterised by low electromagnetic energy fraction and a large ratio of track-p_{T} to calorimeter-E_{T}. On the other hand, low-momentum muons which are absorbed in the calorimeter and overlap with energy deposits from other sources have, inversely, large electromagnetic fraction and low track-p_{T} to calorimeter E_{T} ratio. Using simple cuts on these two variables helps reducing the muon contamination to negligible level. The efficiency of muon veto in true hadronic τ decays is better than 96% whereas the reduction of muons misidentified as τ's is about 40%.

4.1.4 Trigger-level τ reconstruction and identification in Run 1

The technical limitations of the trigger system impose differences of the trigger-level τ reconstruction and identification with respect to offline procedures. At L1 there is nei-



(a) (b)

Figure 4.14: Distributions of f_{HT} (left) and f_{EM} (right) for $Z/\gamma^* \to \tau\tau$ signal and $Z \to ee$ background events [123].

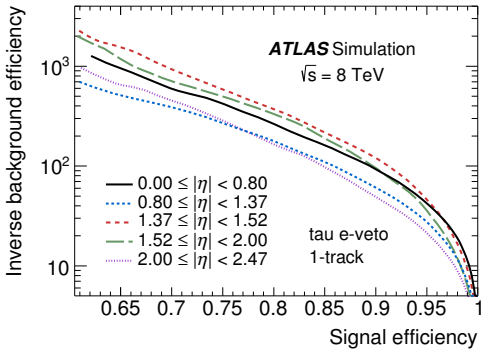


Figure 4.15: Inverse background versus signal efficiency for the electron veto algorithm [123].

ther any track information nor full-granularity calorimeter data available. Topological-clustering algorithm cannot be used at L2 due to time limitations. Finally, at EF the full τ reconstruction and identification is available with the exception of π^0 reconstruction.

At L1 the τ candidates are built from calorimeter energy deposits. Two regions are defined for each candidate containing both the EM and hadronic calorimeters, the core region and the isolation region around the core. The granularity of the trigger towers used is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and a square of 2×2 towers is used to define the core region. The candidate E_{T} at L1 is also measured in this region using two most-energetic neighbouring central towers in the EM calorimeter and all core-region towers in the hadronic calorimeter. The EM isolation energy is obtained from towers between the 0.2×0.2 and 0.4×0.4 regions around the core and is required to be lower than 4 GeV for the lowest-E_{T} threshold at L1.

The RoIs from L1 are used further at L2 as seeds for both calorimeter- and tracking-based observables. Calorimeter observables are computed from cells with pile-up and noise subtracted. The candidate E_{T} is calculated again (with better precision with respect to L1) in the $\Delta R < 0.2$ region around the energy-weighted centre of the τ candidate. Tracking-

based observables are obtained from a fast-tracking algorithm [130] using Pixel and SCT subsystems. With the absence of vertex information at this stage, the rejection of pile-up tracks is performed by placing requirements on the Δz_0 variable which measures longitudinal distance between the track under consideration and highest-p_{T} track in the RoI. High Δz_0 values are typically related to pile-up tracks. In the last step, identification variables, similar to those used in the offline identification, are computed: f_{cent}, f_{track}, R_{track}, with all the limitations present at this stage (EM-level calibration only, no sophisticated pile-up suppression, expanded cone size $\Delta R < 0.4$ for f_{cent} denominator). The L2 τ identification uses cuts on the chosen variables providing inverse background efficiency of about 10 with the efficiency with respect to offline *medium*-identified τ's of about 90%.

The reconstruction and identification of τ leptons at EF level is similar to that of offline version. First, the topological clustering of calorimeter energy deposits is performed within RoIs followed by EF tracking algorithm. With this information available the τ candidate four momentum and identification variables are computed with the exception of π^0-related variables - π^0 reconstruction algorithm was not available at the time when EF trigger algorithms were implemented during Run-1 data taking. Additionally, no pile-up correction is available at this level and no vertex information is present. Therefore, similarly to L2, the Δz_0 requirements are obtained with respect to the leading track and the Δd_0 is computed with respect to the vertex found in the RoI and the cut is set to $\Delta d_0 < 2$ mm. The aforementioned identification variables are used as inputs to BDT algorithms performing suppression of QCD jets separately for single- and three-prong candidates. The working points are defined to obtain 85% and 80% efficiency versus offline *medium*-identified τ's for single-prong and three-prong cases, respectively. With this choice the inverse background efficiency is around 200.

4.1.5 Measurements of τ reconstruction and identification performance in Run-1 data

Experimental tests of the reconstruction and identification algorithms are required to confirm the validity of their implementation and our understanding of the methods used. They are performed in data samples enriched in $Z \to \tau\tau$ events selected with tag-and-probe method. It is based on selecting the events triggered with lepton trigger (in signal process resulting from leptonic decay of one of the τ's) called tag, and containing a hadronic τ decay (probe). Background contamination would always be present in such sample, so its magnitude has to be carefully determined. This measurement was performed with 2012 ATLAS data collected at $\sqrt{s} = 8$ TeV with total integrated luminosity corresponding to 20.3 fb^{-1}.

The measurement is performed with a variable which has a high signal-background separation power. It is chosen as the sum of the number of tracks in the core region and of outer tracks in the $0.2 < \Delta R < 0.6$ region if they meet the requirement of $D^{\mathrm{outer}} = \min([p_{\mathrm{T}}^{\mathrm{core}}/p_{\mathrm{T}}^{\mathrm{outer}}] \cdot \Delta R(\mathrm{core}, \mathrm{outer})) < 4$ where $p_{\mathrm{T}}^{\mathrm{core}}$ is the transverse momentum of any of the core tracks, and the $\Delta R(\mathrm{core}, \mathrm{outer})$ is the angular distance between the candidate outer track and any of the core tracks. This requirement reduces the contribution of pile-up events and allows the signal track-multiplicity spectrum to retain the structure of core-track spectrum. It was chosen to maximise the signal-background separation. For multi-jet background the track multiplicity has larger values due to the inclusion of wider-cone tracks.

Event selection

The $Z/\gamma^* \to \tau\tau$ events with τ_{lep}-τ_{had} final state used in this measurement are triggered with single-electron or single-muon trigger requiring one isolated electron or muon with transverse momentum of $p_{\text{T}} > 24$ GeV. The offline requirements on any of these leptons raise the transverse-momentum threshold to 26 GeV (to avoid trigger-inefficiency regions) and additional track and calorimeter isolation are applied. The electrons are also required to be located within the pseudorapidity region of $|\eta| < 2.47$ with the exception of $1.37 < |\eta| < 1.52$ calorimeter barrel-endcap transition region and to pass *medium* identification criteria. Muons on the other hand have to be contained within $|\eta| < 2.4$. Only events with exactly one such lepton are accepted for further study and the flavour of the lepton determines analysis *channel*. The candidates for hadronic τ decays are required to have $p_{\text{T}} > 15$ GeV, $|\eta| < 2.5$, exactly one or three associated core-region tracks, electric charge of ± 1 and no geometrical overlap with muons of $p_{\text{T}} > 4$ GeV or electrons with $p_{\text{T}} > 15$ GeV passing *medium* or *loose* electron identification, depending on η. In addition, electron- and muon-veto is applied to single-prong τ candidates. A very loose requirement on the τ identification is made by placing a cut on the BDT score retaining 99% of the $Z/\gamma^* \to \tau\tau$ signal events while suppressing a significant amount of jets. The lepton tag and τ-candidate probe are required to have opposite signs (OS) of electric charges. Additional suppression of $Z \to \ell\ell$ and $W \to \ell\nu$ can be achieved by applying requirements related to the topology of events. The invariant mass of the reconstructed τ and tag lepton (usually referred to as *visible* mass because only the directly-observed part of hadronic τ decay enters its computation) is required to lie within 45 GeV $< m_{\text{vis}}(\tau, \ell) < 80$ GeV for τ-p_{T} below 20 GeV or, for higher τ transverse momenta in the muon channel: 50 GeV $< m_{\text{vis}}(\tau, \ell) < 85$ GeV and 50 GeV $< m_{\text{vis}}(\tau, \ell) < 80$ GeV in the electron channel. These regions enclose signal peaks of this variable. The transverse mass of the lepton+$E_{\text{T}}^{\text{miss}}$ system, computed as $m_{\text{T}} = \sqrt{2p_{\text{T}}^{\ell} E_{\text{T}}^{\text{miss}}(1 - \cos\Delta\phi(\ell, E_{\text{T}}^{\text{miss}}))}$ is required to be below 50 GeV. Further reduction of these processes is obtained by requiring that the approximate direction of neutrino system is contained within the angle spanned by the direction of τ and lepton: $\Sigma\cos\Delta\phi = \cos\Delta\phi(\ell, E_{\text{T}}^{\text{miss}}) + \cos\Delta\phi(\tau, E_{\text{T}}^{\text{miss}}) > -0.15$. The two former requirements suppress the background from $W \to \ell\nu$ (+jets) processes in which the τ candidate can emerge from mis-reconstruction of hadronic activity accompanying the produced W boson, and the lepton from W decay is mistakenly used as the tag lepton, as illustrated in Fig. 4.16.

Estimation of background

The track multiplicity distribution of signal sample is obtained from simulated $Z/\gamma^* \to \tau\tau$ sample using only the reconstructed hadronic τ candidates matched to true generator-level τ's. The contribution from QCD multi-jet processes is described by a single template including both the quark- and gluon-initiated jets. It is estimated using a data-driven technique, starting from a phase-space region enriched with multi-jets. This region uses full signal selection with the exception of requiring that the tag lepton and probe τ candidate have the same electric charge signs (same-sign, SS). To estimate the template shape, the contributions from W+jets and Z+jets processes are subtracted. The template is scaled with a OS/SS transition factor estimated in regions with reversed very-loose τ identification criteria. In addition, templates describing the Z+jets and W+jets contributions are added with the former estimated from a simulated sample and the latter extracted from a W-enriched region with removed m_{T} requirement and inverted $\Sigma\cos\Delta\phi$ cut with respect to

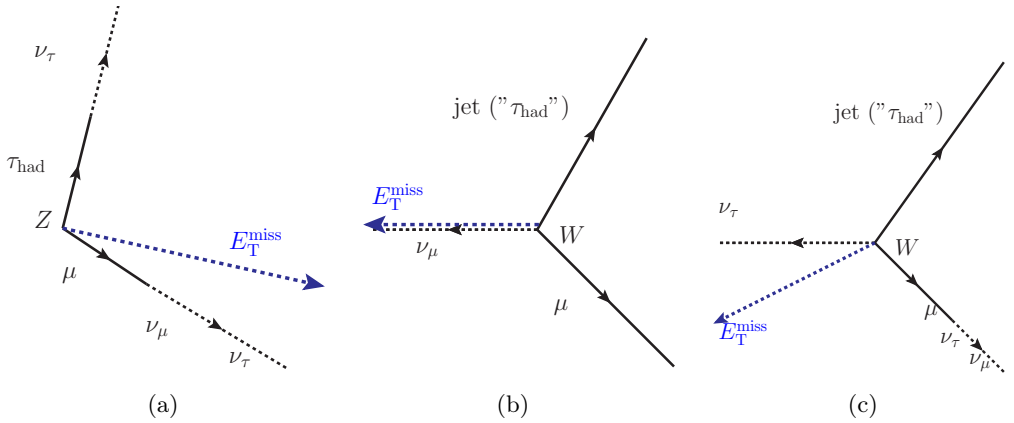


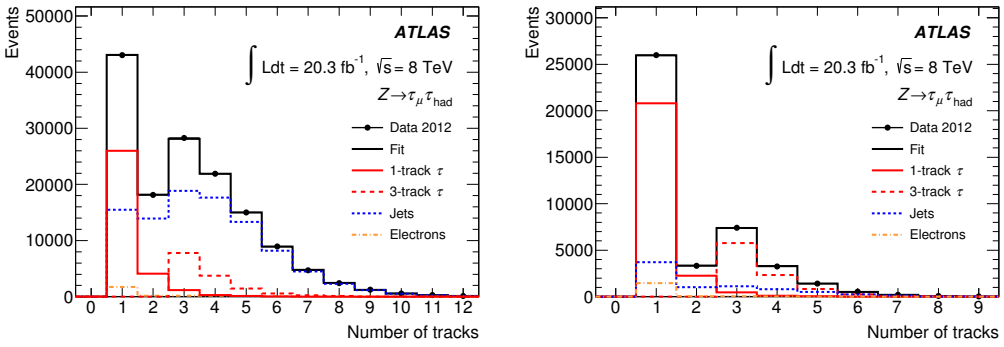
Figure 4.16: Illustrations of the decay product angles in the $Z/\gamma^* \to \tau\tau \to \tau_{\text{lep}}\tau_{\text{had}}$ (left) with hadronic activity on the left side omitted for clarity, $W \to \mu\nu$ (middle) and $W \to \tau\nu \to \mu\nu\nu\nu$ (right).

the signal selection. Simulations are used to estimate the remaining small backgrounds of $Z \to \ell\ell$, $t\bar{t}$ and diboson decays.

Results

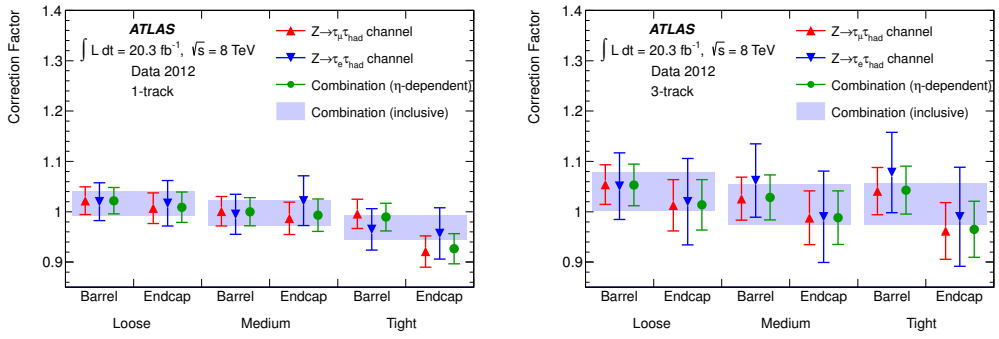
To extract the 1-prong and 3-prong τ efficiencies a fit of the track-multiplicity distribution to the data is performed using the signal and background templates described above. The fit is performed separately in the electron and muon channels and, exclusively, for each of the identification working points, including candidates failing the *loose* identification criteria, those satisfying *loose* but failing *medium* selection, candidates fulfilling *medium* requirements but failing the *tight* and finally τ candidates satisfying *tight* requirements. Figure 4.17 shows examples of these fits for candidates without any identification criteria and for those fulfilling the *medium* requirements. The structure of the track multiplicity spectrum shows characteristic signal peaks at 1 and 3, already before any identification is applied, that become significantly larger after the identification because of high suppression of the background events by the identification algorithm.

The efficiencies of τ identification measured in the data sample show small differences with respect to the values in simulated samples. Correction factors are therefore derived, defined as ratios of the identification efficiency in data to the efficiency in signal simulation for each of the identification working points. To increase the precision of their determination a combination of results from electron and muon channels is performed. Their values are consistent with one with the exception of the *tight* single-prong case. These results are summarised in Fig. 4.18. No significant dependence on τ-p_{T} is observed so the results are subdivided with respect to η region (barrel: $|\eta| < 1.5$, endcap: $1.5 < |\eta| < 2.5$) and track multiplicity only. The uncertainty on the determination of the correction factors are dominated by underlying event and shower model choice in the MC (compared by comparing the default Alpgen+Pythia 6 sample to an auxiliary one produced with Alpgen+Herwig [131, 132]), followed by the uncertainty of jet background estimation. The total uncertainty depends only slightly on the identification working point and totals to about $\pm 2.5\%$ and $\pm 4\%$ in the single-prong and three-prong cases, respectively. The



(a) (b)

Figure 4.17: Results of template fit in the muon channel for the reconstructed τ candidates without any identification criteria (left) and after fulfilling *medium* identification requirements [123].



(a) (b)

Figure 4.18: Correction factors for τ identification efficiencies in simulation obtained from the fit to data as described in the text. Left: single-prong, right: three-prong τ's [123].

results are valid for the τ-p_{T} region between 20 and 100 GeV. Below 20 GeV the uncertainties increase up to 15% and above 100 GeV there are no abundant sources of τ leptons available at the time of writing of this monograph.

Measurement of electron veto performance

To measure the efficiency of electrons, reconstructed as τ candidates, to pass the electron-veto algorithm, a sample of $Z \to ee$ events is selected with high purity using tag-and-probe approach. An electron needs to be identified on one side of the event (tag) and a τ candidate is reconstructed in the other hemisphere (probe). Different combination of electron-veto tightness with various levels of jet discrimination are tested and the efficiencies are extracted in bins of pseudorapidity of the τ candidate. Backgrounds are modelled with simulations with the exception of multi-jet events that are estimated from SS region. Measured differences between the efficiencies in data and MC are, again, parametrized as

correction factors in bins of pseudorapidity. Uncertainties of these factors is also dependent on η and are of the order of 10% for *loose* electron veto and get larger for tighter working points.

τ trigger performance measurement

The efficiency of the τ trigger is also measured with the tag-and-probe method, similar to that used in measurement of offline identification efficiency described earlier. The trigger efficiency is measured with respect to offline-identified τ candidates, therefore the τ-identification criteria are applied in the selection of events. Only muon channel is considered here due to sufficient statistics, lower background contamination and smaller systematic uncertainties compared to the electron channel. In this measurement the dominant systematic uncertainties arise from modelling of background kinematics. The dominant multi-jet and W+jets backgrounds are estimated similarly to the offline efficiency measurement. The considered τ trigger has a calorimetric-isolation requirement and a p_{T} threshold of 11 GeV at L1 and 20 GeV at L2 and EF, up to three tracks and a *medium* identification at EF level. The efficiency turn-on reaches a plateau for $p_{\mathrm{T}} > 35$ GeV and does not depend on the number of primary vertices in the event in neither of the trigger levels. Similarly to the offline case, efficiency correction factors are derived to scale the trigger efficiencies in simulation with respect to those measured in data. The correction factors are compatible with unity with the exception of $p_{\mathrm{T}} < 40$ GeV region where a few-percent deviation is observed. The uncertainty on these factors is at the level of 2% for lower-p_{T} region but increases to about 8% at $p_{\mathrm{T}} > 100$ GeV and is also significant below p_{T} of 30 GeV where the background contamination is highest.

4.1.6 Offline τ-lepton energy calibration in Run 1

The calibration of the τ energy scale (TES) in Run-1 operation of the ATLAS detector is performed in two steps. The first step is based entirely on simulations and is supposed to correct the measured TES to the level of true values. The topological clusters entering τ reconstruction are calibrated with LCW scheme but not optimized for the narrow-cone ($\Delta R < 0.2$) measurements nor for the specific composition of hadrons in the τ jet. The needed correction is derived with simulated τ's from $W \to \tau\nu$, $Z/\gamma^* \to \tau\tau$ and $Z' \to \tau\tau$ decays using only τ candidates with $E_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.4$ matched to true τ leptons, identified at *medium* level and spatially separated ($\Delta R > 0.5$) from other τ candidates. The response is understood as the ratio of reconstructed LCW-scale τ energy to the true visible energy of the candidate. It is evaluated in intervals of true E_{T} and true τ pseudorapidity and, in each interval, fitted with a Gaussian function to determine the mean value. Mean values of the intervals are then fitted with an empirical functional form to determine the response function. On top of it, two additional small corrections are applied, one based on pseudorapidity, to remove bias from poorly-instrumented regions in calorimeters, and another, based on the number of primary vertices in the event – to remove pile-up contribution. The correction factors derived in this step are between 0.97 and 1.07 for single-prong τ candidates and from 0.84 to 1.0 for multi-prong candidates [123].

The second step involves data-driven corrections to the TES and also provides estimates on systematic uncertainties related to the absolute TES and its modelling. Two independent methods can be used. The *deconvolution method* bases on the idea of decomposing each τ candidate into its decay products and combine calorimeter responses

based on τ branching fractions. The response to charged hadrons is obtained from various methods depending on the pseudorapidity and momentum of interest. The response to electromagnetic part was studied with $Z \to ee$ events and is used for neutral pions. The *in-situ* method bases on the measurements of $Z/\gamma^* \to \tau\tau$ events with one of the τ's decaying into muons and the other one into hadrons. They can be used to measure the TES shift between data and simulations by looking at the visible invariant mass distribution in these events. The total uncertainty on the determination of τ energy scale is between 0.9% and 2.4% depending on τ E_{T} and η.

4.1.7 Reconstruction of τ decay products in Run 2

Before the start of LHC Run 2 a new algorithm to improve the overall performance of τ reconstruction, identification and energy calibration, the *Tau Particle Flow* [124] was developed. Its aim is to provide the reconstruction of individual components of hadronic τ decays. The charged hadrons are reconstructed from their tracks in the Inner Detector, whereas neutral pions are reconstructed using energy deposits in the calorimeter. The reconstructed decay products are then used for classification of τ decay mode and to calculate four momentum of the τ candidate. The algorithm operates on τ-candidate objects that were first reconstructed by the baseline reconstruction algorithm, as described in Sect. 4.1.1. Its main focus is to improve the reconstruction of hadronically-decaying τ leptons in the p_{T} range between 15 and 100 GeV which is a typical range of p_{T} of τ leptons produced in the decays of W, Z and SM Higgs bosons.

Neutral pion reconstruction

In the considered transverse momentum range the hadrons from τ decays have lower p_{T}, with peak around 4 GeV. Their typical spatial separation is of the order of $\Delta R \approx 0.07$. The charged hadrons, in addition to creating tracks in the ID, also deposit their energy in calorimeters producing irregular clusters with the usual width of $0.02 < \Delta R < 0.07$. As the π^0's are reconstructed from calorimeter deposits, the main difficulty is to disentangle them from the $\pi^{\pm}$ showers with which they are likely to overlap. The photons from π^0 decays are usually collimated and their angular separation is $0.01 < \Delta R < 0.03$. Therefore, majority of π^0's produces a single cluster in the EM calorimeter. The reconstruction of neutral pions in hadronic τ decays proceeds as follows. First, a π^0 candidate cluster is formed from EM-calorimeter cells found in the core region of the τ decay cone. Its energy is then corrected for the possible contamination from $\pi^{\pm}$ by first estimating its energy deposit in the EM calorimeter by comparing the track momentum with the energy deposited in the hadronic calorimeter, associated to the given track. This contamination is subtracted from the π^0 cluster if their distance is below $\Delta R < 0.04$.

The candidates for π^0's are then subject to identification that helps suppressing the background formed mostly by remnants of $\pi^{\pm}$ contributions or pile-up. Similarly to other places in ATLAS, this procedure also uses BDT algorithm that exploits candidate cluster properties, such as longitudinal (energy fractions in pre-sampler and first two EM-calorimeter layers, cluster depth) and lateral (cluster width, widths of energy deposits in each EM layer, fraction of core energies, etc.) shapes, numbers of cells in each layer and cluster $\eta - \phi$ asymmetry. The identification is optimised in five $|\eta|$ regions to maximise the number of correctly-classified τ decays.

Decay-mode classification

The neutral pion reconstruction and identification procedure described above can be tailored further to provide even better classification of hadronic τ decay modes based on additional information. The π^0 reconstruction mis-classifies about a half of single-prong two-π^0 decays as single-prong one-π^0 type in the simulated $Z/\gamma^* \to \tau\tau$ events, with at least three photons being grouped in a single calorimeter cluster. The fine granularity of the first EM calorimeter layer is used to extract the information on local energy maxima in that layer. It corresponds to the number of photons especially at low π^0 energies. The energy maximum is associated to the π^0 candidate if its transverse energy exceeds 300 - 430 MeV depending on η region. If the E_{T} of the maximum exceeds 10 GeV it is counted as two maxima as this is the case for 95% of such events. Single-prong one-π^0 τ decays are reclassified as single-prong two-π^0 if there are at least three maxima in the first EM layer. This helps recovering 16% of single-prong two-π^0 decays misclassified as single-prong one-π^0 with misclassification rate of 2.5%.

In addition to the number of local energy maxima, another improvement of decay-mode classification can be obtained with the help of kinematics of the τ decay products and π^0 identification scores. A set of three BDTs is trained to perform π^0-counting tests, two for single-prong (0 vs at-least-one π^0 and one vs two-or-more π^0's) and one for three-prong (no π^0's vs at-least-one π^0) τ decays. Inputs to these tests are given by several variables such as π^0 identification scores, their reconstructed masses, number of photons, spatial separation between charged and neutral pion candidates, invariant mass of charged tracks, etc. The resulting performance of the decay-mode classification is summarised in Fig. 4.19. The efficiency for correctly classifying a given decay mode is 74.7%. The estimated uncertainties are negligible.

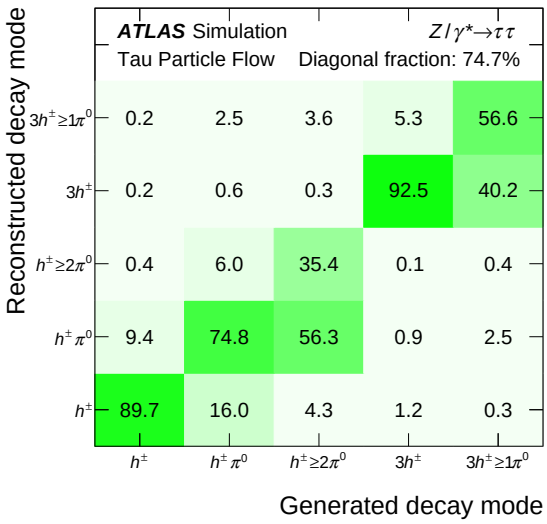


Figure 4.19: Efficiency matrix for decay-mode classification by the Tau Particle Flow algorithm in simulated $Z \to \tau\tau$ events, showing the probability that a given generated τ decay mode will be classified as particular mode. No decays with neutral kaons are considered. The τ leptons have to fulfil basic selection criteria: $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. The statistical uncertainty is negligible [124].

4.1.8 Improvements in τ-lepton identification and energy calibration in Run 2

The τ-identification procedures in Run 2 [125] are based on the same techniques as those of Run 1: BDT-based discriminants are trained separately for single- and three-prong τ candidates using simulated $Z/\gamma^* \to \tau\tau$ (with high-mass Drell-Yan for high-p_{T} τ's) events and di-jet background events extracted from collision data. In Run 2 the identification variables were harmonized between trigger-level and offline-level identification and so was the pile-up correction procedure. Consequently, the variables computed by Run-1 (cluster-based) π^0 identification algorithm (Sec. 4.1.1) were replaced by some of its input variables and by up to two electromagnetic topological clusters to account for the π^0 component. The resulting performance of the algorithm is equivalent to that of the original one with *loose*, *medium* and *tight* working points targeting the efficiencies of 0.6, 0.55 and 0.45 for single-prong τ candidates and 0.5, 0.4 and 0.3 for three-prong τ candidates, respectively.

The baseline energy calibration in Run 2 follows that of Run 1 (Sec. 4.1.6) and works well for high-p_{T} τ candidates but shows a degradation towards lower values of τ p_{T}. A significant improvement in energy resolution at low p_{T} can be achieved using the Tau Particle Flow method, described above (Sec. 4.1.7). It uses constituent-based calculation of τ p_{T} summing four-momenta of charged and neutral pions with the former being measured from the tracking system. The constituent-based calculation is then calibrated in each decay mode to correct for potential biases on π^0 energy. The final calibration of the TES in Run 2 combines the information from Tau Particle Flow algorithm with additional calorimeter and tracking information using boosted regression tree (BRT) [126], a multivariate analysis technique similar to BDT that, instead of providing inputs to decision making, returns the (closer-to) real value of an observable based on many input features. In this implementation [125] the BRT is trained with $Z/\gamma^* \to \tau\tau$ simulated signal events using cluster-shape variables, information on tracks and the input from Tau Particle Flow, namely numbers of neutral pions and relative differences in pion energies. In addition, pile-up-sensitive variables are added to the training to provide pile-up independence of the final response. In addition, an *interpolated* value of τ p_{T}, computed from energy-weighted average of baseline p_{T} and Tau-Particle-Flow-based p_{T} measurements also enters the BRT training. There is a significant improvement in the TES resolution with the BRT-based algorithm in the $p_{\mathrm{T}} < 100$ GeV region, as can be seen in Fig. 4.20.

4.1.9 Run 2 τ performance measurements

Offline identification efficiency

The idea of τ identification performance measurements in Run 2 follows that of Run 1 (Sec. 4.1.5) to large extent and was performed with a data sample of 3.2 fb^{-1}. The tag-and-probe approach is also used here with $Z/\gamma^* \to \tau\tau$ events in the muon channel only, with a very similar event-selection procedure as in Run 1. Multi-jet background is modelled with events in the same-sign region and their contribution scaled with OS to SS ratio obtained in inverse-muon-isolation region. The shape of W+jets contribution is extracted from a W+jets-enriched region in data and scaled to the signal region with MC-based transfer factors. The τ track multiplicity, as defined in Sec. 4.1.5 is used to extract the information on the identification efficiency from data. Simulations are used to determine the shapes of signal and remaining background templates to use in the fit. Similarly to the Run-1 case, efficiency correction factors accounting for efficiency differences between data

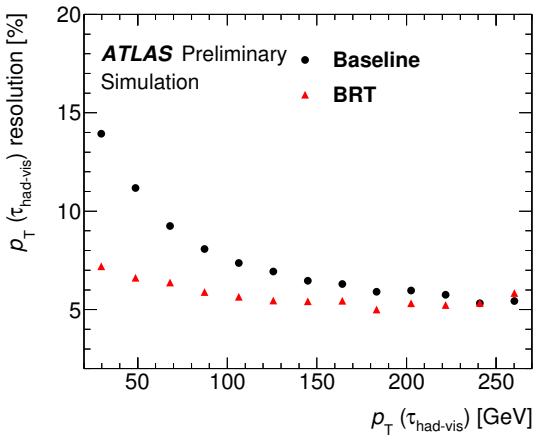


Figure 4.20: The resolution of τ energy scale calibration with the baseline and the BRT-based methods as a function of τ p_{T}. It is defined as the half of the 68% central intervals of the ratio of calibrated p_{T} to the true visible p_{T} value [125].

and simulations are obtained with the fit for each of the working points and for each of τ track multiplicities. They are compatible with unity with the total uncertainties of 5.2% and 5.6% for the *medium* working point in one-prong and three-prong cases, respectively, dominated by systematic uncertainty related to the modelling of signal template [125].

Trigger efficiency

Similar event selection and background estimation procedures and the same dataset as in the above are used for measurement of online τ identification efficiency [125]. It is performed with respect to the offline *medium*-identified τ candidates. The p_{T} thresholds of the considered trigger are 12 GeV at L1 and 25 GeV at HLT. Calorimetric isolation is required at L1 and up to three tracks and *medium* online identification requirements are applied at HLT. The measured efficiency reaches plateau for offline-τ p_{T} of 30 GeV and shows very little dependence on pile-up. It is also compared to that obtained from simulations and correction factors are derived. They are consistent with unity for τ p_{T} above 30 GeV. Their dominant systematic uncertainty is related to the estimation of multi-jet background with a statistically-limited data sample.

τ energy scale

The TES in Run 2 was measured with the in-situ method, similarly to that described in Sec. 4.1.6. It only accounts for the TES shift between data and simulations. In Run 1 the TES shift was determined by the difference in the fitted visible mass peak in data and simulations. An improved method, less prone to statistical fluctuations, was used in the Run-2 measurement. It is based on minimisation of a χ^2 function comparing full shapes of visible mass distributions. The resulting TES shift is compatible with almost no shift for both single-prong and three-prong τ candidates for the baseline TES calibration. The uncertainty of the measured shift factor is 1.4% and 3.2% for single- and three-prong candidates, respectively. The BRT-based energy calibration yields $\alpha = 0.95\% \pm 1.92\%$ and

$\alpha = -3.1\% \pm 1.9\%$ for one- and three-prong candidates, respectively. All these results are dominated by systematic uncertainties related to background modelling.

4.1.10 Summary

The procedures for online and offline τ reconstruction, identification and energy-scale calibration had been established and optimised for the Run-1 data-taking. They had proven to be very successful and robust enough to serve as a basis for the methods used in Run 2. New developments were made and are applied in the analysis of Run-2 data providing similar precision to that of Run 1 despite more difficult running conditions, including higher pile-up and higher rates of multi-jet events at higher collision energies. This allows for continuation and extension of precision studies of Standard Model, including the Higgs sector and also for better sensitivities in searches for BSM phenomena.

4.2 Identification of heavy-flavour jets

The b quark, as any other quark, cannot freely exist for any significant amount of time due to QCD confinement and is subject to hadronisation. As a heavy quark, it undergoes decays into lighter quarks but, as its couplings are suppressed by the off-diagonal CKM matrix elements, the hadrons containing b quarks have a significantly longer lifetime than those composed of lighter quarks. The specific properties of hadrons containing b quarks can be used to identify jets containing b hadrons. These include long lifetime, high mass and decay multiplicities of b hadrons as well as the hard b-quark fragmentation function. There are various algorithms presented in this Section, including those based on decay length significance, full secondary vertex information or impact parameters of the charged tracks [133]. The most powerful observables in the sense of signal-background discrimination are combined using multivariate approach. Additionally, a non-negligible fraction of b jets contains muons from W decays, so an independent algorithm exploiting this feature is also described. The tagging of b jets at trigger level is also briefly described in this Section.

4.2.1 Tagging of b jets in Run 1

The lifetime of hadrons containing b quarks, of the order of 1 ps ($c\tau \approx 300$ μm), is the main feature to help identifying jets containing b quarks. A b hadron with p_{T} of 50 GeV would have a mean length of flight path (before decay) of $\langle l \rangle = \gamma\beta c\tau \approx 2$ mm in the transverse direction. This creates a decay vertex significantly displaced with respect to the primary vertex of hard interaction in pp collision. There are two classes of algorithms exploiting this feature. One is based on determination of longitudinal and transverse impact parameters of tracks and includes JetProb [134] and IP3D [135], while the other, including SV [135] and JetFitter [136] aims for full reconstruction of a displaced secondary vertex. In the end, a combination of different algorithms is performed using multivariate-analysis approach. More details on these algorithms can be found below and in Ref. [133].

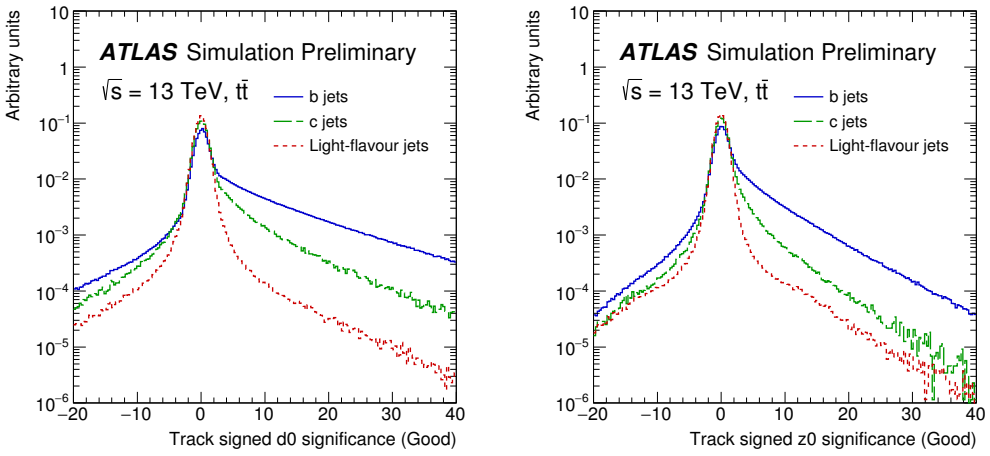
A key part of the lifetime-based b tagging is the reconstruction of the primary vertex [137]. It is the point of reference for both the measurement of impact parameters and displacement of secondary vertices. The precision of its determination depends strongly on the associated track multiplicity - from 200 μm (300 μm in z direction) for two-track vertices to 20 μm (35 μm along z axis) for vertices with 70 associated tracks in minimum-bias events. The precision can be enhanced in events with high-p_{T} objects such as jets or leptons. On the other hand, the number of primary vertices in each collision event was significantly larger than one for most of Run 1, not to mention Run 2. The primary hard-interaction vertex is therefore chosen from all possibilities as the one with the highest sum of p_{T}^2 of the associated tracks.

Another crucial class of ingredients of lifetime-based tagging are, obviously, tracks associated with the jet of interest. The tracks to be used in tagging have to fulfil high-quality criteria that select well-measured tracks and reject those from long-lived particles, material interactions and those with Inner-Detector hits originating from more than one track. At least seven hits in the high-precision silicon tracking detectors are required with at least two in the Pixel sub-detector (and exactly one in the innermost Run-1 layer). The tracks are required to have their p_{T} above 1 GeV and their transverse and longitudinal impact parameters, d_0 and $z_0 \sin\theta$, have to fulfil $|d_0| < 1$ mm and $|z_0 \sin\theta| < 1.5$ mm. These

requirements are very similar to that of τ reconstruction, see Sec. 4.1.1. This selection applies to all impact-parameter based algorithms, whereas the track-selection criteria for secondary-vertex algorithms are loosened to $p_{\mathrm{T}} > 400$ MeV, $|d_0| < 3.5$ mm and no cut on z_0 in the case of the SV algorithm, and $p_{\mathrm{T}} > 500$ MeV, $|d_0| < 7$ mm and $|z_0 \sin\theta| < 10$ mm in the JetFitter case. Both these algorithms require at least one hit in the Pixel detector.

Impact parameter-based algorithms

The impact parameters are computed with respect to the primary hard-interaction vertex. A sign is associated to the impact parameter to further discriminate primary-vertex from b-hadron tracks. The sign is defined as positive if the track intersects the jet axis in front of the primary vertex (or, in other words, on the same side where the jet is located) and negative otherwise. The jet axis is defined by the calorimeter-based direction. If a secondary vertex is found in the jet, the jet direction is updated to the direction of the line between the primary and secondary vertices. The tracks originating from the primary vertex would usually have a random sign distribution whereas the tracks from b and c jets would most often yield a positive impact-parameter sign. This is also the case for decays of e.g. K_S^0 or Λ^0 so an enhancement of light-jet identification as b jets may occur. Fig. 4.21 provides an illustration of the impact parameter distributions for various jet types. It was created using Run-2 simulations but the behaviour is the same also for lower-energy collisions in Run 1.



(a) (b)

Figure 4.21: Transverse (left) and longitudinal (right) impact parameter significances for b (blue), c (green) and light-flavour (red) jets as obtained from the simulated Run-2 $t\bar{t}$ events [138]. Very similar behaviour was observed in Run-1 data.

The JetProb algorithm, previously used in LEP and Tevatron experiments, bases on the impact parameter significance, defined as $S_{d_0} = d_0/\sigma_{d_0}$ where σ_{d_0} is the uncertainty of d_0 measurement. The value of S_{d_0} is compared to a pre-determined resolution function to measure the probability that the given track originates from the primary vertex. The resolution function is obtained from experimental data with the information extracted

from the negative side of signed impact parameter distribution assuming no heavy-flavour contribution in that region. The resulting probability has a uniform distribution for light jets and shows a sharp peak around 0 for b jets. This purely data-driven algorithm was used in the 2011 data taking in ATLAS and for the whole Run 1 in trigger-level b tagging.

A powerful algorithm relying on both the transverse and longitudinal impact parameter significances ($d_0/\sigma_{d_0},\ z_0/\sigma_{z_0}$) and their correlations is the IP3D. It is based on a log-likelihood ratio (LLR) method that compares the values of the measured track significances to the pre-defined two-dimensional probability-density functions obtained from simulations of b and light-flavour jets. A weight for each track is defined by the ratio of probabilities and the jet weight is constructed as a sum of logarithms of individual weights of the associated tracks.

Vertex-based algorithms

A three-dimensional information on a secondary vertex formed by decay products of a b hadron can further help discriminating between b and light-flavour jets. The basic algorithms of the vertex-based type use jet-associated tracks with significant displacement from the primary vertex (three-dimensional impact parameter significance $S_{d_{3D}} > 2$). Vertex candidates are then formed for track pairs with a χ^2 fit. All the vertex candidates that can be associated with material interactions are rejected by checking if the vertex position overlaps with the location of innermost Pixel layers. Vertices from K^0_S and Λ^0 decays as well as photon conversions are suppressed using the invariant mass of charged-particle track four-momenta. All the remaining two-track vertices are then combined into a single inclusive vertex iteratively removing tracks with largest χ^2 contribution to the vertex fit until a predefined threshold is passed.

The simplest vertex-based algorithm in ATLAS is the SV0. It relies on the three-dimensional flight length significance, the distance between primary and secondary vertices divided by its uncertainty. Similarly to the impact parameters, the flight length is also signed with respect to the jet direction. The SV0 yields much smaller mis-tag rate than impact-parameter-based taggers but for a price of secondary-vertex-finding efficiency reduced down to 70%. This is a typical feature of the vertex-based algorithms.

The SV1 on the other hand is a more sophisticated vertex-based b-tagging algorithm that relies on the same secondary-vertex finding procedure but benefits from the application of LLR formalism, like the aforementioned IP3D. It exploits the invariant mass of tracks (assuming they are pions) associated to the vertex (lower for light-flavour jets), the ratio of the sum of energies of these tracks with respect to all tracks associated with the jet of interest (higher in b jets), the number of two-track vertices and the angular distance between the jet direction and the direction of line between the primary and secondary vertices. The efficiency of identifying b jets with this algorithm is between 60% and 70% while it does not exceed 35% for c jets and 15% for lighter-flavour jets.

The JetFitter algorithm employs a different approach and bases on topological structure of decays inside a jet. First, a common line between the primary vertex and the vertices of b and c decays is sought with a Kalman filter together with the positions of these vertices along that line to approximate the b-hadron flight path. The vertices are not merged together in this approach but their number is used as one of the discriminating variables instead. The others include the number of vertices with a single associated track, the total number of tracks in the decay chain, the invariant mass of all tracks attached to the decay chain, the energy fraction of these tracks with respect to all tracks associated with

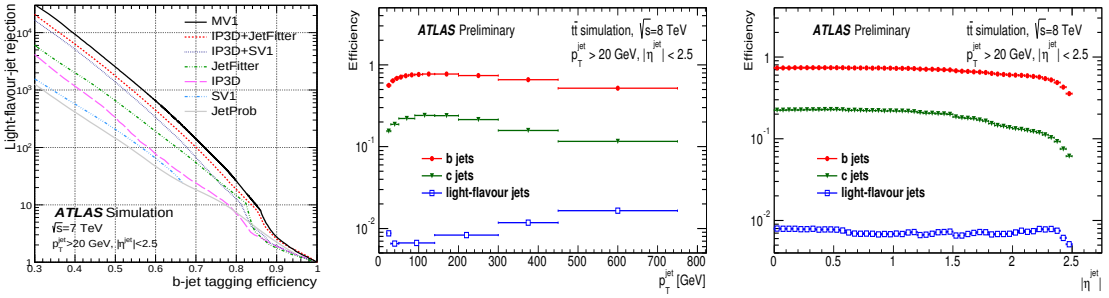
the examined jet and the flight length significance averaged over the individual vertices, weighted with the squared inverse of their decay-length uncertainties. These variables are combined in an Artificial Neural Network (ANN) algorithm [139] together with p_{T} and $|\eta|$ of the examined jets as the distributions of the discriminating variables show dependence on jet kinematics. The final output of this algorithm is a set of b-tagging probabilities for b, c and light-flavour jets that are used to obtain the final discriminant score.

MVA-based combinations

A combination of the secondary-vertex- and impact-parameter-based approaches to b-jet tagging can help extracting the advantages of both of them while minimizing the potential drawbacks. The LLR-based algorithms, SV1 and IP3D can be combined in a very straightforward way, by summing their respective output weights forming the so-called IP3D+SV1 algorithm. Another possibility is to combine the ANN-based JetFitter with IP3D using the output IP3D weight as an additional input variable to the otherwise unchanged JetFitter ANN, constituting the IP3D+JetFitter algorithm. However, it is the MV1 algorithm that was most commonly used in the ATLAS Run-1 analyses. It combines the IP3D, SV1 and the sum of IP3D and JetFitter discriminants into a Multi-Layer Perceptron (MLP) neural network. Not only it benefits from the discrimination power of several algorithms but also it exploits the correlations between them. These correlations are low when considering IP3D tagger with respect to vertex-based algorithms while the output from SV1 and IP3D+JetFitter is correlated but the correlations are different for the different jet categories (b, c and light flavour) under consideration. The MV1 is trained using a simulated samples containing b jets (signal) and light-jets (background) obtained from simulated $t\bar{t}$ decays. In addition, simulated di-jet events with 200 GeV $< p_{\mathrm{T}} <$ 500 GeV are also included in the training. A re-weighting based on p_{T} and $|\eta|$ of jets is used to reduce the potential bias from different kinematic spectra of different jet categories. As usual in the case of multivariate analysis-based identification algorithms, a set of working points is defined - in this case for b-jet selection efficiencies of 60%, 70% and 80%. The resulting simulation-based performance of all the described b-tagging algorithms is shown in Fig. 4.22 (left) together with the efficiency for various jet flavours of the MV1 algorithm (middle & right). Eventually, for the analyses based on the full Run-1 dataset a version of MV1, named MV1c, trained against a mixture of light and c-jets to enhance the c-jet rejection was used.

b-tagging performance measurements

The efficiency of the MV1 tagger was measured in data with the 2012 data sample corresponding to 20.2 fb^{-1} collected at $\sqrt{s} = 8$ TeV [140, 141]. Similarly to the case of τ-lepton reconstruction (see Sec. 4.1.5 and 4.1.9) correction factors are derived to bring the simulated efficiencies for b, c and light-flavour jets closer to the values measured in collision data. The efficiency of b-tagging is measured in a data sample dominated with $t\bar{t}$ events with two oppositely-charged leptons in the final state. The resulting correction factors were obtained for transverse momenta intervals between 20 and 300 GeV and are in most cases consistent with one. Their uncertainties range from 2% in the p_{T} region around 100 GeV up to 7-8% at 70% working point for lowest and highest p_{T} values with the dominant contributions from jet energy-scale and background normalization uncertainties. They are presented in Fig. 4.23 (left).

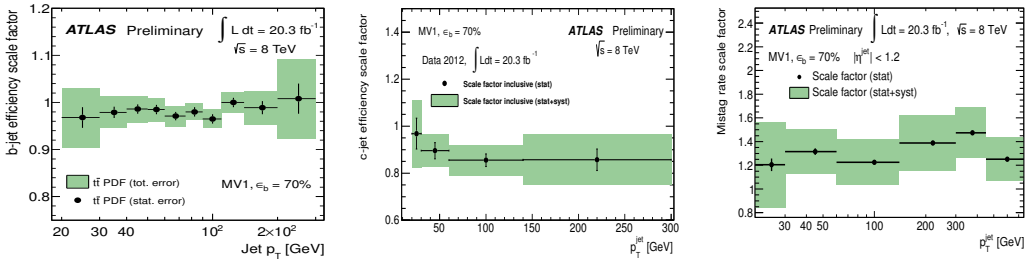


(a) (b) (c)

Figure 4.22: Rejection of light jets (inverse efficiency) versus b-tagging efficiency for various Run-1 b-tagging algorithms (left) [133], Efficiency of the MV1 tagger to select b, c and light-flavour jets at 70% working point with respect to their p_{T} (middle) and $|\eta|$ (right) [140].

The efficiency for c-jet tagging was measured using D^{*+} mesons produced in b decays and decaying into $D^0(\rightarrow K^-\pi^+)\pi^+$. Pairs of charged tracks are first combined to form a D^0 meson in its expected mass region and then with a track with an opposite charge with respect to the kaon candidate. The D^{*+} is associated with the jet under consideration if their angular distance is $\Delta R < 0.3$. The resulting correction factors are usually around 0.9 with uncertainties between 8% and 15% for the 70% efficiency working point. The c-jet efficiency scale factors are shown in the middle panel of Fig. 4.23.

Estimation of mis-tag rate (a fraction of light-flavour jets mistakenly tagged as b jets) is performed with a data sample selected by inverting the sign of the flight length significance of tracks for the impact parameter-based taggers or the decay length significances for the vertex-based algorithms as for the prompt tracks the distributions of these parameters is expected to be symmetric. The correction factors for mis-tag rates are usually slightly above unity and their uncertainties range between 15% and 43%. The results for central ($|\eta| < 1.2$) jets are shown in the right panel of Fig. 4.23. The same conclusions hold for $|\eta| > 1.2$ jets.



(a) (b) (c)

Figure 4.23: MV1 tagger performance correction factors at 70% working point. Correction factors for b-jet efficiency (left) [141], correction factors for c-jet identification efficiency (middle) [140] and MV1 mis-tag rate correction factors (right) [140].

Muon-based b tagging

In about 20% of b jets a decay of b hadron into muons may occur, either directly, or via $b \to c \to \mu$ cascade decays. Therefore the efficiency of muon-based b-tagging algorithms is intrinsincly lower than that of lifetime-based ones, however the muon-based taggers do not use any lifetime information, what makes them complementary to the other techniques and subject to different sources of systematic uncertainty. The Soft Muon Tagger (SMT) [133] uses combined muons associated with the primary vertex to reject particles from pile-up interactions. The muon is then associated with the closest jet if their angular separation is $\Delta R < 0.5$. Additional requirements on the muon are $p_{\mathrm{T}} > 4$ GeV, $|d_0| < 3$ mm and $|z_0 \cdot \sin\theta| < 3$ mm. The SMT also uses the χ^2 value of the ID-to-MS track matching in the muon reconstruction. Larger values typically occur for in-flight decays of light charged mesons ($\pi^{\pm}$, $K^{\pm}$) that significantly contribute to the sample of jets with associated muons.

The performance of the SMT algorithm was estimated using simulated samples of $t\bar{t}$ and inclusive-jet events. The average efficiencies for b- and c-jets tagging by the SMT algorithm are found to be 11.1% and 4.4%, respectively, with lower efficiencies for lower values of jet p_{T}. The mis-tag rates are very low, below 0.5%.

Trigger-level b-tagging implementation

The triggering of b jets starts with L1 selection of jet candidates. The tracking information on which the b-tagging algorithms base becomes available at L2 and EF where an online version of the aforementioned JetProb algorithm is implemented and applied to L1 jets. Three working points, *loose*, *medium* and *tight* are defined corresponding to the efficiencies of 70%, 55% and 40% with respect to the true offline b jets, measured on a simulated $t\bar{t}$ sample. A refined jet reconstruction has also been implemented at L2 and EF to bring the online jet p_{T} closer to that of the offline measurement. Various combinations of jet-p_{T} thresholds, b-tagging working points and jet multiplicities are used to maximise the acceptance for different analyses.

4.2.2 Improvements in b tagging in Run 2

Several changes were implemented in the ATLAS b-tagging infrastructure before and during Run 2 [138, 142]. First of all, the installation of the Insertable B-Layer [100] had a significant impact on the decay parameter resolution. In addition, improvements were made in the tracking algorithms including the use of ANN algorithm for clustering of hits in the Pixel detector [143], resulting in improved handling of hits shared between multiple tracks which often happens in the core of high-p_{T} jets [144].

Impact parameter-based tagging

Two impact-parameter-based b-tagging algorithms were used during Run 2. The IP3D, already described above (Sec. 4.2.1), and IP2D that applies the same design principles but uses only transverse impact parameter significance. The latter is more robust against pile-up as it does not take into account the longitudinal impact parameter significance which is typically large not only for b-jet tracks but also for tracks in jets emerging from additional interaction vertices. In both these algorithms the LLR-based approach is applied. In addition, several refinements of the IP-based algorithms have been introduced during Run 2. The required number of hits in the Pixel detector was changed from two to one to increase

the selection efficiency for high-p_{T} b jets where a significant fraction of b hadrons decay after the IBL and the first Pixel layer. A sizeable gain in light-flavour jet rejection can be achieved by ignoring tracks from conversions and decays of K_S and Λ (about 15% at 77% working point). In addition, not only the simulated $t\bar{t}$ events are used for the reference probability-density functions in the LLR approach but also $Z' \to t\bar{t}$.

Vertex-based tagging

In Run 2 a single secondary-vertex tagger, the SV, is in use. It bases on the same principles as the Run-1 algorithms in terms of selecting tracks to reconstruct two-track vertices and iteratively combining them into a single vertex with outliers removed at each iteration. Improvements with respect to Run 1 include the imposing of the limit on the number of tracks associated with the secondary vertex of 25 most energetic tracks, to reduce the number of fake vertices often occurring in high-p_{T} (above 300 GeV) jets due to high multiplicity of jet-fragmentation tracks. An additional track cleaning is applied to jets in the $|\eta| > 1.5$ region to mitigate the effects of higher amount of detector material and resulting worse track parameter resolution and increased rate of hadronic interactions. Tracks in that region are required to have eight, instead of seven hits in the silicon sub-detectors. A reduction of pile-up effects is achieved by removing low $S_{d_0} < 2$ and high $S_{z_0} > 6$ tracks from the selection. In addition, a set of requirements is also imposed on the selected two-track vertices to further reduce fake contribution. These include a pixel hit pattern corresponding to the vertex radius for both tracks, an invariant mass smaller than 6 GeV and not corresponding to the masses of K_S^0 and Λ^0. The resulting secondary-vertex reconstruction efficiency in the simulated $t\bar{t}$ sample is similar to that of Run 1 with a small increase of b-tagging efficiency and nearly the same efficiencies for c and light-flavour jets as in Run 1.

The JetFitter algorithm is also present among the Run-2 ATLAS b taggers. No improvements were implemented in this algorithm with respect to its Run-1 version.

MVA taggers

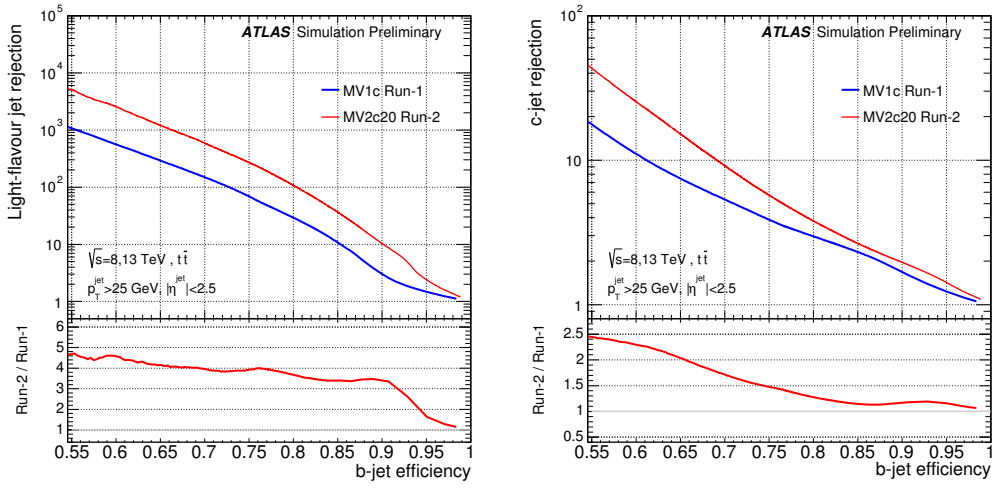
The idea of combining the outputs from the various b-tagging algorithms with various features, already developed for Run-1 analyses has also been in place for Run-2-based studies. The MV2 [142] algorithm combines the output from the three basic algorithms using a BDT technique rather than ANN (Run 1). The training was performed using a simulated $t\bar{t}$ sample in three variants. The MV2c20, used in the first part of Run 2, was trained with a background sample composed in 80% of light-flavour jets with a 20%-contribution of c jets. Starting from 2016, the MV2c10 variant, with 93% (7%)[3] of light-flavour (c) jets in the training signal sample, has been the main ATLAS b-tagging algorithm. The inclusion of c jets in the training results in a slight reduction of light-flavour jet rejection rates but also in a strong increase of c-jet rejection with respect to the MV2c00 variant, trained with the light-flavour jets in the background sample. Four operating points were chosen to provide a defined efficiency of 60%, 70%, 77% and 85% for the b jets in the $t\bar{t}$ sample. The corresponding rejection rates are listed in Table 4.1. The overall gain in the b-tagging performance in Run 2 with respect to that of Run 1 can

[3]It was decided to keep the MV2cXX nomenclature where initially the XX stated for the percentage of c jets in the background training sample despite the fact it does not reflect the actual c-jet percentage from 2016 onwards.

b-jet efficiency	c-jet rejection	light-jet rejection
60%	34	1538
70%	12	381
77%	6	134
85%	3.1	33

Table 4.1: MV2c10 b-tagging algorithm operating points with the corresponding rejection rates estimated using simulated $t\bar{t}$ events [138].

be seen in Fig. 4.24. At the 70%-efficiency working point the improvement in rejection of light jets by a factor of about 4 was observed, whereas the c-jet rejection was improved by a factor of 1.5-2. In other words, a relative 10% gain in b-jet efficiency was achieved. This large improvement is partly thanks to the installation of IBL that results in a much better impact parameter resolution for tracks with p_{T} of 5 - 10 GeV [100]. It affects the low and medium jet-p_{T} region. In higher p_{T} the improvements are due to the application of new algorithms.



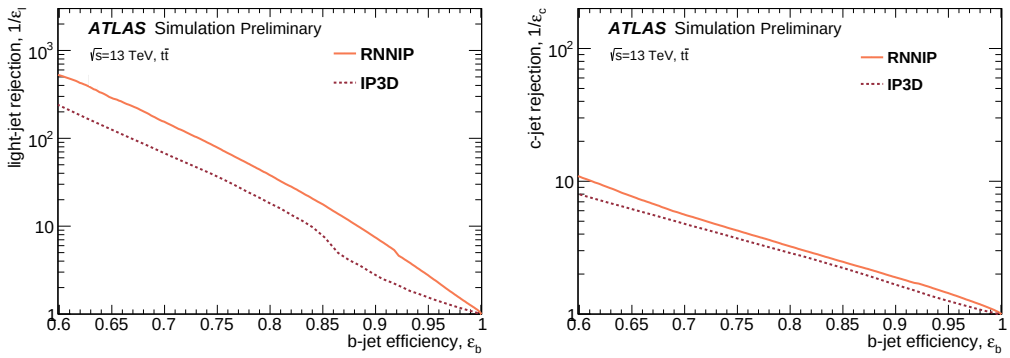
(a) (b)

Figure 4.24: Light (left) and c-jet (right) rejection versus the b-tagging efficiency for the Run-1 MV1 (blue) and Run-2 MV2c20 (red) algorithms obtained in the simulated $t\bar{t}$ sample. [142]. No public plots for the MV2c10 algorithm were available at the time of writing of this monograph but the improvement with respect to Run-1 performance is similar.

The MV2c20 algorithm was also adapted to work in the HLT [102]. Operating points analogous to *loose*, *medium* and *tight* were defined and their light-jet rejection is at the level of that of Run-1 offline b-tagging.

Improvements for 2017-2018 data taking

In the second part of Run 2 a Recurrent Neural Network [145] tagger, RNNIP [146], was introduced to improve the impact parameter-based tagging by exploiting track-to-track correlations. They are present by definition when several charged particles are produced in the secondary vertex with large impact parameters: when one of the tracks is found to have a large impact parameter, the others are also likely to show similar properties. No such correlation should on the other hand exist in case of no displaced secondary vertex as in light-flavour jets. The baseline IP3D algorithm uses the LLR approach with likelihood templates computed assuming the properties of each track in a jet are independent, due to large sample statistics needed to compute such templates. This however prevents the ability to fully model the properties of b jets. The Recurrent Neural Network approach can be used to overcome this limitation by directly learning sequential dependencies between the input variables. The input set of variables in this case includes the significances of both the transverse and longitudinal impact parameters, the fraction of p_{T} carried by the given track with respect to the p_{T} of the jet under consideration, the angular distance between the track and the jet axis and the track score obtained from the IP2D and IP3D algorithms. The RNNIP returns b-tagging probabilities for b, c, light-flavour and τ jets. They are further used as inputs to the combined taggers but can also serve as a base for constructing the discriminant score to use the RNNIP as a standalone tagger. Not surprisingly, it outperforms the basic IP3D, as can be observed in Fig. 4.25.



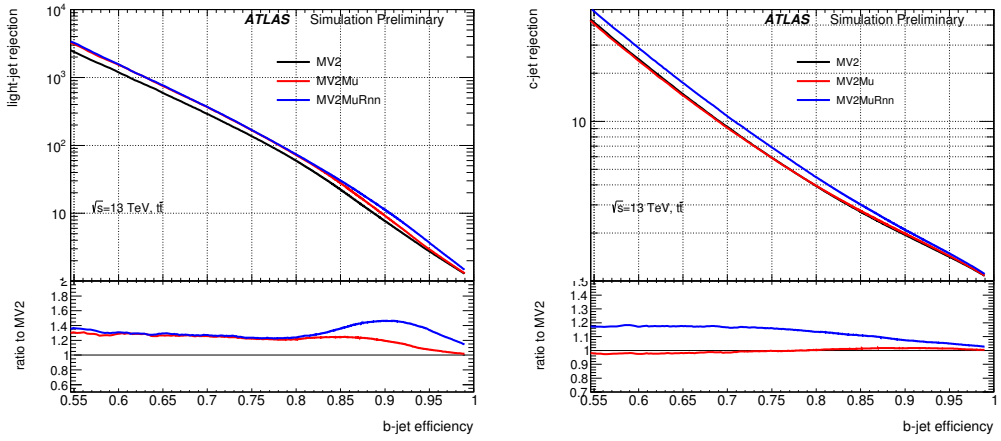
(a) (b)

Figure 4.25: Light (left) and c-jet (right) rejection versus the b-tagging efficiency for the Run-2 IP3D (dashed) and RNNIP (solid) algorithms obtained in the simulated $t\bar{t}$ sample [147].

In addition to the introduction of RNNIP, a new version of the SMT algorithm [147] was also developed and implemented for the 2017-2018 data taking. In addition to using combined muons and associating them to nearby calorimeter jets the new version benefits from an improved rejection of background from prompt muons randomly associated with light jets, muons from in-flight decays of pions and kaons and the punch-through of energetic hadrons that travel through the calorimeter and reach the MS. This improved rejection is obtained using three discriminating variables: ΔR (typically smaller for c jets), d_0 (larger on absolute value for b jets) and a muon transverse momentum relative to the jet axis, usually larger for b jets. A set of another three variables quantifying muon-track

quality is also constructed to exploit the information on potential kink along the muon track (that may indicate the pion or kaon decay), the imbalance in momentum measured in the ID and MS (also helpful to suppress the in-flight hadron decays) and the ratio of q/p (charge over momentum) between ID and MS. The full set of the six listed variables is used as an input to a BDT algorithm. An efficiency of 85% for b jets containing an SMT-candidate muon with 15% mis-tag rate is achieved with the cut on the BDT discriminant. For a b-tagging efficiency of 10% a mis-tag rate of 0.02% is obtained.

Improvements were also made in the area of MVA-based b tagging. A hybrid training sample consisting not only of the simulated prompt $t\bar{t}$ events but also of the $Z' \to t\bar{t}$ is used to cover a wider range the jet transverse momentum. Moreover, the availability of the new and refined basic taggers, new variants of the MV2 algorithms were introduced, including the reference option (unchanged with respect to 2016), the MV2MU that uses SMT in addition and the full option using RNNIP and SMT with the standard MV2 input variables (MV2MURNN). The resulting performance of the new variants is presented together with that of basic MV2 in Fig. 4.26. An improvement is observed for the full variant in all cases. In addition, a new MVA-based tagging algorithm, the DL1, was



(a) (b)

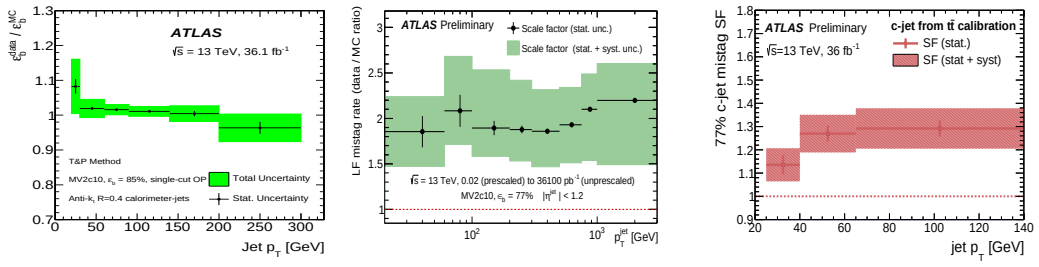
Figure 4.26: Light (left) and c-jet (right) rejection versus the b-tagging efficiency for the Run-2 basic MV2 (black), MV2MU (red) and MV2MURNN (blue) algorithm variants obtained in the simulated $t\bar{t}$ sample [147].

introduced. It is based on the Deep Neural Networks (DNN) [148, 149]. Its input consists of the same set of variables as that of MV2 tagger with the full SMT observable set instead of just its output. Its performance was found to be very similar to that of MV2. The use of DNN-based approach was however motivated by the possible further development of b-tagging algorithms and can, in principle, provide a better architecture to exploit input correlations than the BDT.

Run-2 b-tagging performance measurements

The efficiency of the MV2C10 algorithm was studied in Run-2 data using the 36 fb^{-1} dataset collected in 2015-2016 using $t\bar{t}$ events selected with the tag-and-probe method .

The data-to-simulation correction factors obtained from this measurement are found to be consistent with unity with the uncertainty of below 10% for the 70%-efficiency working point, see Fig. 4.27 left [150]. The efficiency for light-flavour jet b tagging (mis-tag rate) was studied [151] and the resulting data-to-simulation correction factors are around 2 with the uncertainty of 0.5 at 77%-efficiency working point (Fig. 4.27 middle). Finally, a measurement of b-tagging efficiency in c jets was also performed with the same dataset [152]. The correction factors obtained are between 1.1 and 1.3 with the uncertainty below 0.1, see Fig. 4.27 right.



(a) (b) (c)

Figure 4.27: Left: b-tagging efficiency correction factors of the MV2c10 tagger at 70% efficiency working point [150]. Middle: correction factors for the light-flavour mis-tag rates from the MV2c10 algorithm for central jets at 77% working point [151]. Right: correction factors for the c-jet (mis)tagging efficiency of the MV2c10 algorithm at 77% working point [152].

4.2.3 Summary

The development of methods for identifying jets containing b hadrons before and during Run 1 was successful and established a basis for the improvements in the Run-2 b-tagging. These were obviously helped by the installation of the IBL sub-detector, but also a large amount of work was put into the development of new and refined algorithms that can fully benefit from all the available information by exploiting machine-learning techniques. They provide optimal use of all the available information including not only the discriminating variables themselves but also correlations between them. This is particularly important for analyses searching for rare signals in the vast amounts of ATLAS data.

Chapter 5

Standard Model processes with third-generation-fermion final states

The measurements in the Standard Model sector have been an important part of the ATLAS physics programme since the beginning of the LHC operation. Increasing amounts of collision data together with detector upgrades and developments of more robust and performant algorithms to reconstruct and identify physics objects all contribute to the increase of precision of such measurements. This enhanced precision brings a better understanding of the sub-atomic world of particles but also provides increasingly better description of processes constituting experimental backgrounds for searches for Beyond-Standard-Model phenomena. From 2012 onwards also the measurements in the Higgs sector can be legitimately included to the area of Standard Model physics.

5.1 Measurement of τ-lepton polarisation in $Z/\gamma^* \to \tau\tau$ decays

The polarisation of τ lepton, P_τ, is defined as the relative difference between the cross-section for positive (σ_+) and negative (σ_-) helicity τ lepton production:

$$P_\tau = \frac{\sigma_+ - \sigma_-}{\sigma_+ + \sigma_-} \tag{5.1}$$

for the τ^- lepton. It measures the degree of parity violation in the τ lepton production process providing information on its Lorentz structure. The left-handed (right-handed) chiral states coincide with the negative (positive) helicity states in the relativistic limit of massless τ's assumed here and are experimentally accessible as the τ leptons decay inside the ATLAS detector volume and the kinematic distributions of decay products are sensitive to the spin of the τ lepton. These distributions are equivalent for the left-handed τ^- and right-handed τ^+ leptons and vice-versa due to nearly-exact CP invariance in τ decays, therefore in the following only one of the states is mentioned at a time.

The first measurements of τ lepton polarisation produced in $Z/\gamma^* \to \tau\tau$ decays were performed at LEP in electron-positron annihilation events at the Z-boson mass pole. The most precise value of the average τ polarisation was obtained in the combination of LEP results and is presented as the τ production asymmetry, A_τ, a quantity very close to P_τ with reversed sign and a small $\mathcal{O}(0.005)$ correction from Z and photon propagators interference. The asymmetry value was $A_\tau = 0.144 \pm 0.004$ [153].

The ATLAS measurement described in this Section is documented in Ref. [154] in detail. It is the first measurement of τ polarisation in $Z/\gamma^* \to \tau\tau$ decays performed at a hadron collider. It provides a complementary constraint on the polarisation of τ leptons in decays of Z/γ^* produced in a qqZ vertex where quark-electroweak couplings are involved. It is performed in the $66 < m_{Z/\gamma^*} < 116$ GeV mass range as the contributions from slightly above and below the Z-mass pole cannot be accurately separated due to experimental limitations. The contributions to P_τ away from the Z mass peak depend on the interference between the Z boson and photon-mediated amplitudes. These contributions have approximately opposite effect above and below the mass peak. Because of that and of the dominant cross-section of the on-pole contribution the on-pole polarisation is very close to that of the wider invariant mass region. Theoretical predictions from the Alpgen [128] generator interfaced with Pythia 6 [129] parton shower and hadronisation model and Tauola [155, 156] for the τ decays is $P_\tau = -0.152 \pm 0.002$.

An earlier ATLAS measurement of τ polarisation involved $W \to \tau\nu$ decays [157] and was performed with the 2010 dataset corresponding to 24 pb^{-1}. The idea of measuring τ polarisation with a template fit to a polarisation-sensitive observable was inherited from that study, yet new techniques were implemented to estimate systematic uncertainties and to model significant background contributions in the measurement described here. This analysis may serve as a foundation for further measurements in decays of the Higgs boson into $\tau\tau$ pairs. The polarisation itself can serve to distinguish the vector Z boson from the scalar H and its potential pseudo-scalar counterpart A. In addition, in the $H^+ \to \tau\nu$ analysis, described in Sec. 6.3 a polarisation-sensitive variable optimised for the $Z/\gamma^* \to \tau\tau$ polarisation studies, is used to enhance the search sensitivity.

5.1.1 Data and simulated samples

This measurement is based on the 2012 proton – proton dataset collected by ATLAS with the centre-of-mass energy of $\sqrt{s} = 8$ TeV and corresponding to the integrated luminosity of 20.2 fb^{-1} after event quality criteria are fulfilled. Triggers accepting a single light lepton are used to select candidate events with p_{T} thresholds of 24 GeV for both electrons and muons and additional isolation requirements, aiming to efficiently select $Z/\gamma^* \to \tau\tau$ events with lepton-hadron final state, following the tag-and-probe approach as in many other analyses with τ leptons (see Sec. 4.1.5 for examples).

Simulated samples are used to model the expected contribution from signal and some of the background processes. Signal $Z/\gamma^* \to \tau\tau$ (+jets) events are generated with the Alpgen generator interfaced with Pythia 6 for fragmentation and hadronisation and the decays of τ leptons are simulated with the Tauola program. The helicities of τ leptons, not stored after Tauola simulation were re-simulated using the TauSpinner software package [158–160] based on the kinematic properties of τ decays. Auxiliary signal samples for studies of systematic uncertainties were generated with Pythia 8 and Alpgen+Herwig/Jimmy generators. The Alpgen+Pythia 6+Tauola setup was also used to produce background samples of the decays of W and Z bosons into τ's and light leptons with accompanying jets. Production of $t\bar{t}$ pairs was made with the Powheg+Pythia 6 generator setup.

5.1.2 Event selection

The selection of actual signal event sample is based on the properties of both the leptonic (tag) and the hadronic (probe) leg of the $Z/\gamma^* \to \tau\tau$ decay and on the expected signal event topology. The lepton is required to have $p_{\mathrm{T}} > 26$ GeV and the transverse momentum of tracks in a $\Delta R < 0.4$ cone around it should not exceed 6% of the lepton momentum. Similarly, transverse energy deposited in the $\Delta R < 0.2$ cone around the lepton axis (and not associated to the lepton itself) must be below 6% of its transverse energy. The electrons are required to pass *tight* identification criteria whereas for muons only the existence of ID+MS-combined track is required. Events are categorised into channels by the lepton flavour (further referred to as muon channel and electron channel). The hadronically-decaying τ leptons must fulfil the $p_{\mathrm{T}} > 20$ GeV requirement together with the *medium* identification against jets, a *medium* electron veto and muon veto requirements. Only single-prong τ candidates are retained for this analysis. Exactly one such τ candidate and exactly one light lepton fulfilling the requirements listed above and of opposite electric charges are required. In addition, the visible mass of the τ+lepton system is required to be contained within $40 < m_{\mathrm{vis}} < 85$ GeV.

To further suppress the background contribution from W+jets events a set of requirements on the event topology is imposed. The criteria on transverse mass, m_{T} (defined in Sec. 4.1.5) and sum of azimuthal angles, $\Sigma\Delta\phi = \Delta\phi(\tau, E_{\mathrm{T}}^{\mathrm{miss}}) + \Delta\phi(\mathrm{lepton}, E_{\mathrm{T}}^{\mathrm{miss}})$, are required to satisfy $m_{\mathrm{T}} < 30$ GeV and $\Sigma\Delta\phi < 0.15$, respectively. The procedure for selecting events described above defines the signal region of this analysis. Additional regions for background estimations are defined in Section 5.1.4.

Some of the event selection requirements exhibit different acceptances for the two τ helicity states. The lepton-p_{T} requirement efficiency for selecting left-handed τ^- states is almost twice as efficient as for right-handed τ^- states as the hadronic-τ spin state is correlated to that of leptonic τ due to the angular momentum conservation. A partial

compensation of this efficiency bias is observed for the τ-p_{T} and transverse mass requirements.

5.1.3 Polarisation-sensitive observable

The simplest τ decay channel, exhibiting the highest sensitivity to the τ polarization is the $\tau^{\pm} \to \pi^{\pm}\nu$ final state. In that case, the neutrino (which is always left-handed) is preferably emitted opposite the spin orientation of the τ to conserve angular momentum. The angle θ between the τ direction of flight and charged pion in the τ rest frame is the primary observable sensitive to τ polarization. Unfortunately that angle cannot be directly measured since only the hadronic decay products are observed in the ATLAS detector and neither the τ direction nor the direction of the ν from the τ decay can be determined.

The $\tau^{\pm} \to \rho^{\pm}\nu$ channel offers the kinematic simplicity of a two-body decay, like the $\tau^{\pm} \to \pi^{\pm}\nu$ channel, but the dynamics are more complicated. This is because the ρ is a vector particle and thus has two possible helicity configurations. It can either be longitudinally (helicity equals 0) or transversely polarised (helicity equals ± 1). The sensitivity of the angle θ is lower in this mode, compared to the pion decay due to the mixing of longitudinally and transversely polarized vector states. This loss of sensitivity can be compensated by analysing the subsequent decay $\rho^{\pm} \to \pi^{\pm}\pi^{0}$, whose energy and angular distributions depend on the ρ helicity. For decays into a ρ meson, a variable sensitive to τ helicity is the angle between the direction of flight of the vector meson and the charged pion in the rest frame of the vector meson. The energy sharing between the charged and the neutral pion depends on this angle. Thus for these decays, a quantity referred to as the *charged asymmetry* is defined as:

$$\Upsilon_{\mathrm{th}} = \frac{E_{\pi^{\pm}} - E_{\pi^{0}}}{E_{\pi^{\pm}} + E_{\pi^{0}}}. \tag{5.2}$$

This variable was used in the ATLAS studies on τ polarisation in the decays of W bosons [157].

Experimentally, the energy associated with the charged pion can be obtained from the transverse momentum of the single track associated with the τ candidate. The energy ascribed to the neutral pion(s) is calculated from the difference between the τ lepton visible p_{T} and the track p_{T} of the τ candidate. By the time of performing this analysis the TauParticleFlow algorithm (see Sec. 4.1.7) was not available and the cluster-based π^{0} reconstruction (see Sec. 4.1.1) was not fully optimised. Therefore neither the classification of τ decay modes nor the reconstruction of π^{0} kinematics could help increasing the precision of this measurement. The charged asymmetry Υ_{th} (Equation 5.2) is approximated as follows:

$$\Upsilon = \frac{E_{\mathrm{T}}^{\pi^{\pm}} - E_{\mathrm{T}}^{\pi^{0}}}{E_{\mathrm{T}}^{\tau}} = 2\frac{p_{\mathrm{T}}^{\mathrm{track}}}{E_{\mathrm{T}}^{\tau}} - 1. \tag{5.3}$$

Figure 5.1 shows the distributions of Υ for inclusive 1-prong τ candidates after the full event selection.

5.1.4 Estimation of background

The background processes for the measurements based on the tag-and-probe approach in $Z/\gamma^{*} \to \tau\tau$ decays have already been listed in Sec. 4.1.5. The same classes of processes have to be handled in this measurement and similar techniques are used here.

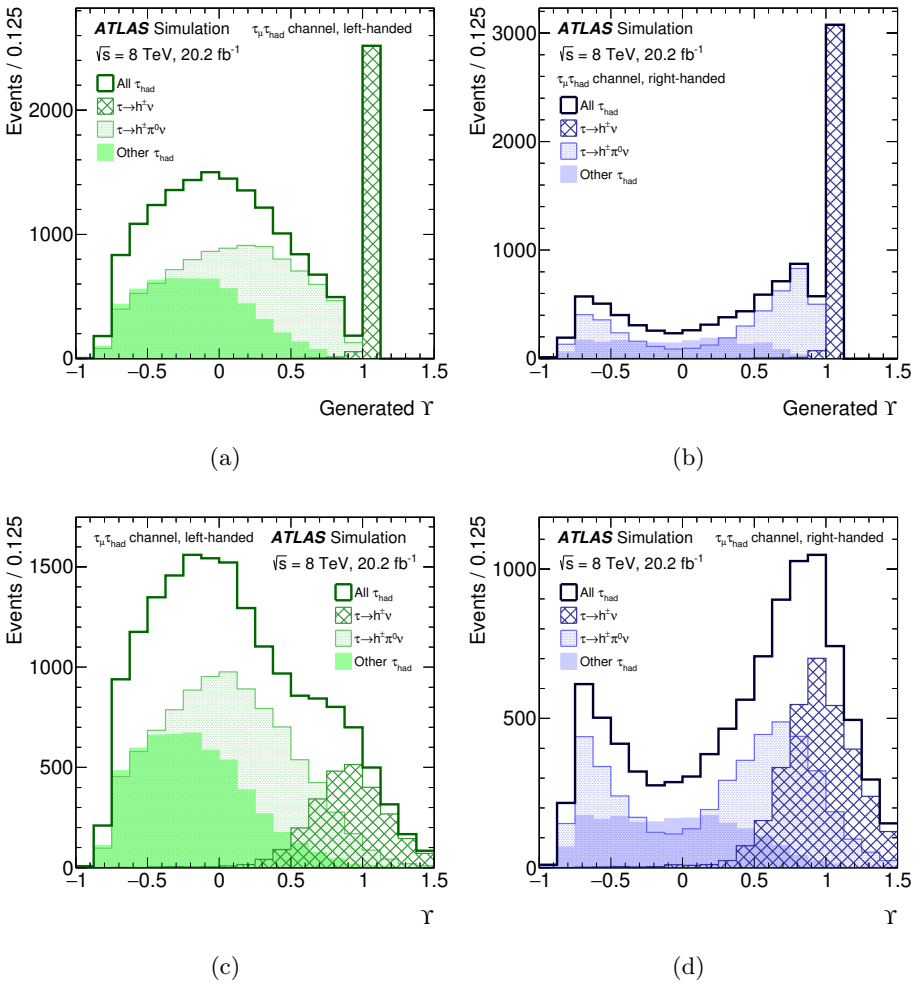


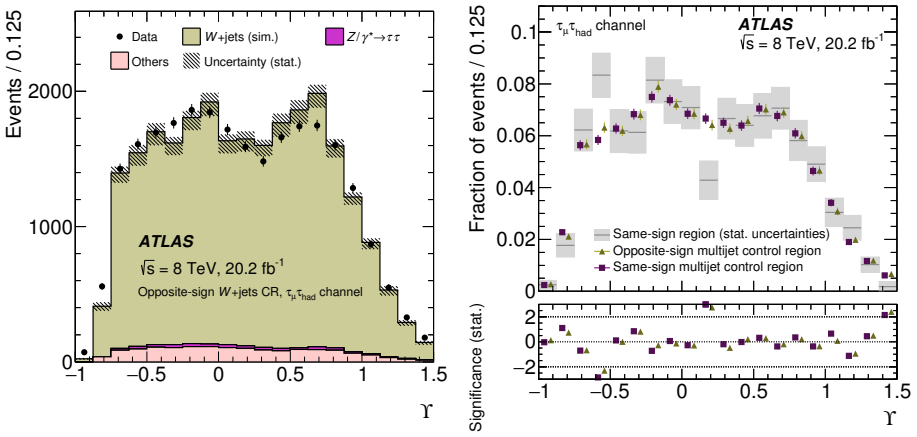
Figure 5.1: Distributions of the charged asymmetry observable for the left-handed (left column) and right-handed (right column) τ leptons from the Z/γ^* decays in the muon channel after full event selection. Top row shows the stable-particle level distributions and bottom row - those obtained at reconstructed-detector level. Distributions for individual decay modes are overlaid for illustration [154]. Distributions in the electron channel exhibit a similar behaviour.

The contributions from two major backgrounds, the QCD multi-jets and the $W \to \ell\nu(+\text{jets})$ are estimated in a data-driven way.

The shape of the Υ distribution in W+jets events is obtained from a control region defined as the signal region but with inverted $\Sigma\Delta\phi$ requirement and the transverse mass required to be $m_{\mathrm{T}} > 70$ GeV. A simulation-based subtraction of the contributions from other processes is performed with the exception of multi-jet background that is negligible in that region. This includes the contribution from the $Z/\gamma^* \to \tau\tau$ signal. A potential measurement bias from a non-SM value of τ polarisation in these decays was checked (and is included in the estimation of systematic uncertainties) but was found to be negligible. The normalisation of the W+jets contribution is obtained by scaling the control-region shape

by the ratio of W+jets event yields in the signal and the control regions, extracted from simulation. Fig. 5.2 (left) shows the shape of Υ distribution in the W-control region for the muon channel with the simulation-based contributions from various processes overlaid.

The multi-jet background shape is extracted from a control region defined by selection criteria following those of the signal region with the exception of requiring the τ+lepton candidate pair to carry the same signs (SS) of electric charges. This shape is then scaled by an OS/SS scale factor obtained from another control regions, with inverted lepton-isolation requirements. In these control regions, the contributions from other processes are subtracted. The W+jets contribution is estimated as described above and all other contributions are extracted from simulations, including that of $Z/\gamma^* \to \tau\tau$ signal, which, similarly to the W+jets case, has a negligible impact on the overall measurement precision. The value of the OS/SS scale factor was estimated to be 1.05 (1.12) in the electron (muon) channel. It was verified that the shapes of Υ distributions in these control regions are in agreement, as can be seen in Fig. 5.2 (right).



(a) (b)

Figure 5.2: Left: the Υ distribution in the W-control region with the contributions from individual processes estimated from simulations. Right: Multi-jet Υ distribution shapes for all three control regions used in the multi-jet background estimation (see text) in the muon channel. Distributions in the electron channel exhibit a similar behaviour [154].

Minor background contributions from the $Z \to \ell\ell$ and $t\bar{t}$ processes are estimated directly from simulations. They amount to about 5% (2%) of the total event yield in the electron (muon) channel.

5.1.5 Systematic uncertainties

This analysis relies on the prediction of the overall shapes and normalisations of the signal and the various background Υ templates.

The dominant contribution to the systematic uncertainty results from uncertainties in signal modelling, related to the choice of event generator. These are estimated by comparing the default ALPGEN signal sample to the auxiliary ones produced with PYTHIA 8 and POWHEG generators. Event weights with respect to the kinematics of the Z boson and to the Υ spectra are obtained this way for various τ decay modes. The resulting

uncertainties arise from the uncertainties on modelling the τ pseudorapidity and the Υ itself. In addition, also the use of TauSpinner to split the simulated $Z/\gamma^* \to \tau\tau$ samples into left- and right-handed components may affect the Υ distributions of these individual components. This is estimated by varying the TauSpinner input parameters, such as QCD factorisation and renormalisation scales, α_S and the PDFs.

Some of the variables used in the identification of hadronic τ decays exhibit strong correlations with the Υ variable. Mis-modelling of those variables affects the shape of Υ. This effect was studied by comparing the identification variables in simulations to those obtained from data in W- and $t\bar{t}$-enriched samples. In addition, the energy response to the hadronic τ decays can also affect the modelling of Υ shape. The TES uncertainty, described in Sec. 4.1.6 and in Ref. [123] is in this analysis split into hadronic and electromagnetic parts. The remaining uncertainties have minor effect on the final result.

5.1.6 Fit and results

The τ polarisation is estimated with an extended binned maximum-likelihood fit to the Υ variable. The fit is performed simultaneously in the signal region and in the SS background control region in both the electron and muon channels. Three signal Υ histograms are used in the fit to describe the left- and right-handed component as well as a minor component of events migrating from outside of the $66 < m_{Z/\gamma^*} < 116$ GeV mass window region. The left- and right-handed templates are scaled to the full $Z/\gamma^* \to \tau\tau$ cross-section and the polarisation value, P_τ, is extracted as the relative normalisation parameter. The simulated Υ distributions of the $Z \to \ell\ell$ and $t\bar{t}$ and the data-driven W+jets template histograms, normalised as described in Sec. 5.1.4, are also used in the fit. The multi-jet background is extracted from a simultaneous fit in the SS region with nuisance parameters common for these two regions for each bin and channel. The OS/SS scale factors are fixed in the fit. The fit model contains nuisance parameters accounting for the limited statistics of the MC signal samples and for systematic variations of the shapes and normalizations of the histogram templates used in the fit. Individual sources of uncertainties are considered uncorrelated. The statistical uncertainty related to the statistics of Monte Carlo signal samples is estimated with a variation of the Barlow-Beeston method [161]. In this analysis, each bin in each region is given a single Poisson-constrained nuisance parameter associated with the total Monte Carlo estimate and the total statistical uncertainty in that bin.

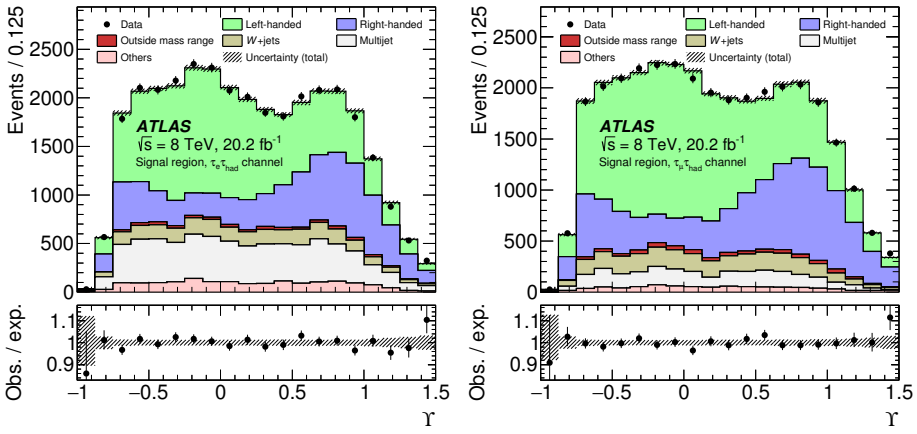
The resulting values of τ polarisation in the $Z/\gamma^* \to \tau\tau$ decays in the mass region of $66 < m_{Z/\gamma^*} < 116$ GeV, extracted in the electron and muon channels, as well as their combination, are presented in Table 5.1 and are in agreement with the Standard Model predictions. The distributions of the Υ shapes after the combined fit are depicted in Fig. 5.3. The impact of different systematic-uncertainty sources is summarised in Table 5.2.

5.1.7 Summary

The ATLAS measurement of the τ polarisation in the decays of Z/γ^* boson was performed with the 2012 dataset and yielded no deviation from the Standard model prediction. The resulting value of the τ polarisation is $P_\tau = -0.14 \pm 0.02(\text{stat}) \pm 0.04(\text{syst})$. No corresponding result was published by the CMS collaboration, however a PhD thesis on

Channel	P_τ in the mass window
Electron	-0.20 ± 0.02 (stat) $\pm$ 0.05 (syst)
Muon	-0.13 ± 0.02 (stat) $\pm$ 0.05 (syst)
Combination	-0.14 ± 0.02 (stat) $\pm$ 0.04 (syst)

Table 5.1: Measured τ polarisation values and overall uncertainties in the mass region of $66 < m_{Z/\gamma^*} < 116$ GeV [154].



(a) (b)

Figure 5.3: Post-fit Υ distributions in the electron (left) and muon (right) channels [154].

similar analysis, with SM-compatible results and similar precision to that of the described measurement was presented in 2016 [162].

Source of uncertainty	Contribution to σ_{P_τ}
Modelling of signal process	0.026
τ identification	0.020
MC statistical	0.016
Signal sample splitting	0.015
TES	0.015
Multi-jet estimate	0.013
Other	0.009
Total systematic uncertainty	0.040
Statistical uncertainty	0.015

Table 5.2: Impact of the individual sources of uncertainty on the polarisation uncertainty σ_{P_τ} in the combined fits [154].

5.2 Higgs boson decays into pairs of τ leptons

This section provides a concise description of the ATLAS measurements involving the decays of Standard Model Higgs boson into a pair of τ leptons, starting from its first evidence for the $H \to \tau\tau$ coupling in Run 1 [37], through the tests of CP invariance in these decays [163] and the Run-2 cross-section measurement for this process [164].

5.2.1 Evidence for the $H \to \tau\tau$ Yukawa coupling

The coupling of the Higgs boson to the electroweak vector bosons was confirmed with the Higgs discovery measurements [30, 31]. The fermion mass-generation mechanism is however of a similar importance in the SM as well as the proportionality of its strength to the masses of fermions [165]. The most favourable channel for the first searches of such couplings is the $H \to \tau\tau$ due to a better signal-to-background ratio than that of $H \to b\bar{b}$.

Samples

The analysis providing the first evidence for $H \to \tau\tau$ couplings in ATLAS [37] was performed with the full Run-1 dataset, collected at centre-of-mass energies of $\sqrt{s} = 7$ and 8 TeV with the integrated luminosities corresponding to 4.5 fb^{-1} and 20.2 fb^{-1}, respectively. Different final-state signatures are expected for the different considered Higgs production processes, the ggF, VBF and VH, in addition to three types of final states involving the three channels of τ-pair decays: τ_{lep}-τ_{lep}, τ_{lep}-τ_{had} and τ_{had}-τ_{had}. This implied using both the single-lepton, di-lepton and di-τ triggers for event selection.

Simulated MC samples were used to model the contribution from the expected signal as well as of several of the background processes. The Powheg+Pythia 8 setup was used to produce both the ggF and VBF signal, whereas Pythia 8 was used to model the VH production. Background MC samples involving electroweak boson decays were produced with the Powheg+Pythia 8 setup similarly to the contribution from top-quark decays. Diboson backgrounds were estimated with the help of Herwig, Alpgen+Herwig and gg2WW [166]+Herwig setups. A dominant irreducible background from the $Z/\gamma^* \to \tau\tau$ process was modelled using so-called *embedded* hybrid sample, in which the $Z/\gamma^* \to \mu\mu$ decays are extracted from real collision data with high purity and the isolated high-energy muons are then replaced with τ decays simulated with the Tauola generator. This minimises the dependence on simulation in estimation of this particular background contribution.

Selection and categorisation of events

Events are first selected into three mutually-exclusive categories resembling τ-pair decay channels. The τ_{lep}-τ_{lep} channel selection requires exactly two isolated opposite-sign leptons with the one with leading p_{T} above 20-25 GeV and the sub-leading one passing p_{T} threshold of 10-15 GeV, depending on the data-taking period and the applied trigger thresholds. Electrons are required to fulfil *medium* identification criteria (see Sec. 3.4). In this category, events with hadronic τ candidates are rejected. In order to suppress background contributions from Z/γ^* and quarkonia decays, requirements on the visible di-lepton invariant mass and angular separation of leptons are imposed. Additionally, $E_{\text{T}}^{\text{miss}} > 20$ GeV selection is made in the different-flavour ($e\mu$) case whereas for the same-flavour events a special type of high-p_{T} object-based missing transverse energy, $E_{\text{T}}^{\text{miss HPTO}} >$

40 GeV is used to suppress Z/γ^* decays. Additionally, the momentum carried by the undetected neutrino system is approximated with the *collinear approximation* [167] and the visible fractions of τ momenta carried by leptons are required to be between 0.1 and 1.0. A potential overlap with the $H \to WW^* \to \ell\nu\ell\nu$ analysis is removed by requiring the $\tau\tau$ invariant mass from collinear approximation to exceed 66 GeV. Events containing b-tagged jets with p_{T} above 25 GeV are also rejected.

The τ_{lep}-τ_{had} channel selection criteria require the presence of exactly one isolated lepton (with *medium* identification in case of electron) with p_{T} above 22-26 GeV, depending on the data-taking period, and exactly one isolated 20-GeV τ candidate fulfilling *medium* identification criteria, with its electric charge sign opposite to that of the lepton. Suppression of W+jets background is achieved by requiring the transverse mass to be $m_{\mathrm{T}} < 70$ GeV. Events with b-tagged 30-GeV jets are rejected.

Two opposite-sign τ candidates with 35 and 25-GeV p_{T} are required for the τ_{had}-τ_{had} channel. One of them has to fulfil the *tight* identification criteria while for the other the *medium* requirements are imposed. Events with light leptons are rejected. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be above 20 GeV and its direction should lie between the two τ candidates in the azimuthal angle. Additional cuts on angular separation on the two τ candidates are imposed to further suppress multi-jet background.

Two exclusive event categories are defined to cover the expected signal topologies. The *VBF* category is characterised by the presence of two additional high-p_{T} jets with large separation in η. Therefore, the requirement on $\Delta\eta(j1, j2)$ between the two leading-p_{T} jets is imposed. In the τ_{had}-τ_{had} channel a requirement on the visible invariant mass of the $\tau\tau$ system is required to be above 40 GeV to suppress low-mass Z/γ^* events. This category does include a small admixture of ggF and VH events. The *boosted* category on the other hand targets ggF events. The Higgs boson candidates are required to exhibit a p_{T} over 100 GeV, calculated from the vectorial sum visible τ decay products and the $E_{\mathrm{T}}^{\mathrm{miss}}$. The VBF-category selection criteria have to be failed. In addition, in the τ_{lep}-τ_{lep} channel a presence of an additional 40-GeV jet is required. This category includes smaller contributions of the VBF and VH events.

Discriminating variable

The discriminating variable to extract the potential signal contribution from data is defined by the output from the BDT algorithm trained separately in each event category and each channel with individual sets of input variables. The most important variables used as BDT inputs are the following. The di-τ invariant mass, $m_{\tau\tau}^{\mathrm{MMC}}$ is reconstructed with the missing mass calculator (MMC) [168] by solving an under-constrained system of equations for six to eight unknowns, using constraints from the x and y components of $E_{\mathrm{T}}^{\mathrm{miss}}$ and visible masses of τ's followed by a scan over the two components of $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and the yet undetermined variables. Each scan point is weighted by its probability according to the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution and the τ decay topology. The estimated $\tau\tau$ mass is the most probable value of the scan points. Additionally, the BDT uses angular separation between τ's, pseudorapidity separation between two leading jets (VBF category), total p_{T} of the visible τ decay products, leading jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, relative angular $E_{\mathrm{T}}^{\mathrm{miss}}$ direction and other quantities describing event topologies.

Estimation of background

The dominant background process, common to all analysis channels, is the $Z/\gamma^* \to \tau\tau$ decay. As said in Sec. 5.2.1, it is estimated with the help of the embedded sample in which the $Z \to \mu\mu$ decays are selected from collision data and only the τ decays are simulated, replacing the reconstructed muons but retaining their kinematics. Estimation of this background contribution directly from $Z/\gamma^* \to \tau\tau$ Monte Carlo samples would result in much larger associated systematic uncertainties. On the other hand, selecting a sufficiently pure signal-free control region for this process was not possible. The embedding procedure was thoroughly tested and validated using both data and simulations.

Other background sources in the τ_{lep}-τ_{lep} channel are treated inclusively, using a common control region for W+jets, $t\bar{t}$ and multi-jets, constructed by inverting lepton isolation criteria compared to the signal regions selection. The $Z \to \mu\mu$ and $Z \to ee$ contributions are obtained from simulations and subtracted. The normalisation of the inclusive background is estimated with a fit to p_{T} spectrum of the sub-leading lepton.

The so-called *fake-factor* method was used to estimate the background contributions from W/Z+jets, multi-jets and semi-leptonic $t\bar{t}$ decays in the τ_{lep}-τ_{had} channel. The fake factor is defined as the ratio of the number of jets identified as *medium* τ candidates to those identified as *loose* but not *medium*. They are derived from jet-enriched samples of W+jets, $t\bar{t}$ and Z+jets backgrounds obtained by selecting $m_{\text{T}} > 70$ GeV region, reversing b-tagging requirement and selecting Z-mass peak region in the di-lepton invariant mass spectrum, respectively. A multi-jet sample is selected by setting the τ identification criteria to *loose* in the event selection. These p_{T}-dependent factors are then applied to events selected with the signal-selection criteria with the exception of requiring the hadronic τ candidate to fulfil the *loose* selection criteria but to fail *medium* requirements.

The multi-jet background in the τ_{had}-τ_{had} channel is extracted from control regions defined as VBF or boosted signal regions but with inverted τ isolation and opposite-sign requirements. The normalisation is determined with a simultaneous fit of the multi-jet and $Z/\gamma^* \to \tau\tau$ contributions to the $\tau - \tau$ pseudorapidity difference distribution. In this channel other background contributions are estimated from simulations.

Systematic uncertainties

The BDT input variables and therefore the BDT output are all subject to systematic uncertainties, as well as the numbers of expected signal and background events are. The dominant sources of the uncertainties in all channels are related to the jet energy scale determination. The uncertainty on background estimation is sub-dominant. Significant systematic contributions arise from the uncertainties on τ energy scale and identification efficiency. Uncertainties on the modelling of Higgs branching fractions, the model of parton shower and the choice of PDFs for the simulations are also found to be non-negligible.

Statistical analysis and results

The BDT output distributions in all the six analysis categories for both the $\sqrt{s} = 7$ and 8 TeV are used as the final discriminating variable in the simultaneous maximum-likelihood fit to extract the signal strength parameter $\mu_{H\to\tau\tau}$, where $\mu_{H\to\tau\tau} = 0.0$ corresponds to the absence of signal and $\mu_{H\to\tau\tau} = 1.0$ corresponds to the presence of Higgs boson signal with cross-section equal to that of the SM predictions. The information from control regions is

included to constrain background normalisations. Systematic uncertainties are taken into account by means of nuisance parameters and affect the final results. Figure 5.4 shows the BDT distributions in all the categories with all the normalisations and nuisance parameters adjusted with the fit for the $\sqrt{s} = 8$ TeV dataset. A good agreement between the data and the model with Standard-Model Higgs boson at the mass of 125 GeV is found. The signal

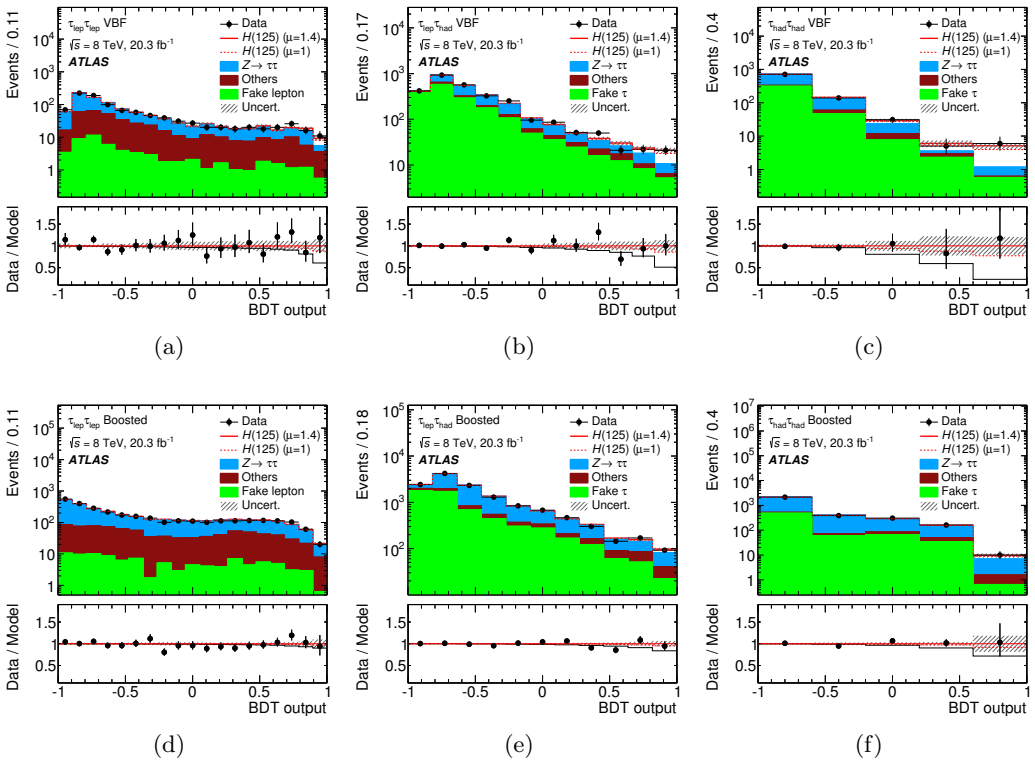


Figure 5.4: BDT distributions in the $\sqrt{s} = 8$ TeV dataset for the VBF (top row) and boosted (bottom row) categories in the τ_{lep}-τ_{lep} (left column), τ_{lep}-τ_{had} (middle column) and τ_{had}-τ_{had} (right column) channels. The Higgs boson signal with mass of 125 GeV is also shown with signal strength of 1.0 (dashed line) and 1.4 (solid line) times the SM expectations. All normalisations are obtained with the final fit. In the ratio panels the dashed and solid lines represent SM expectations for signal strength of 1.0 and 0.0, respectively [37].

strength obtained in this analysis at the Higgs boson mass of 125 GeV is [37]:

$$\mu_{H\to\tau\tau} = 1.43^{+0.27}_{-0.26}(\text{stat})^{+0.33}_{-0.27}(\text{syst}). \tag{5.4}$$

Summary

Evidence for decays of the Higgs boson into τ leptons was found with the analysis of full Run-1 ATLAS dataset using the combination of all three possible decay channels of $\tau\tau$ pairs. The excess above the expected background from other SM processes was found with an observed (expected) significance of 4.5 (3.4) standard deviations. This is consistent with the SM expectations for the Higgs boson with mass of 125 GeV and with the predicted Yukawa coupling strength.

5.2.2 CP-invariance tests in $H \to \tau\tau$ decays

This analysis [163] constitutes a first direct test of CP invariance in VBF Higgs boson production. It employs a CP-odd *Optimal Observable* [169–171] and is based on the same triggers, event selection, background estimation and systematic uncertainties determination procedures as the $H \to \tau\tau$ evidence analysis, described in Section 5.2.1 and in Ref. [37]. The τ_{lep}-τ_{lep} and τ_{lep}-τ_{had} channels are used. Also, the simulated Monte Carlo samples are retained from that analysis. The 2012 ATLAS dataset corresponding to 20.2 fb^{-1} collected at $\sqrt{s} = 8$ TeV is used here.

The Optimal Observable

The matrix element for the VBF Higgs boson production, including a non-SM CP-violating term can be written as $\mathcal{M} = \mathcal{M}_{\text{SM}} + \tilde{d} \cdot \mathcal{M}_{CP-\text{odd}}$. The cross-section (matrix-element squared) is then $|\mathcal{M}|^2 = |\mathcal{M}_{\text{SM}}|^2 + \tilde{d} \cdot 2Re(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{CP-\text{odd}}) + \tilde{d}^2|\mathcal{M}_{CP-\text{odd}}|^2$. Only the middle, interference term, proportional to $\tilde{d}$ is CP-odd and is a possible CP-violation source in the Higgs sector. The optimal observable has to be proportional to that term and is defined as $\mathcal{OO} = \frac{2Re(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{CP-\text{odd}})}{|\mathcal{M}_{\text{SM}}|^2}$. A non-zero value of the $\mathcal{OO}$ would be a clear sign of CP-violation in the Higgs sector.

The values for the leading-order matrix elements for these calculations are extracted using the HAWK program [172–174] with the four-momenta of Higgs and tagging jets as an input. Figure 5.5 shows the distributions of the observable for three values of CP-violating $\tilde{d}$. They were obtained with the MADGRAPH5_AMC@NLO [175] event generator-based re-weighting of the nominal signal sample to the non-SM $\tilde{d}$ values.

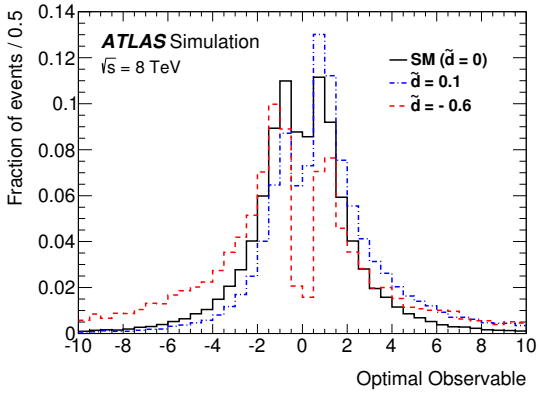


Figure 5.5: Distribution of the Optimal Observable for three $\tilde{d}$ values [163].

The analysis

As said, the VBF selection from the $H \to \tau\tau$ evidence analysis is used here with the same BDTs. In this analysis the final signal selection requires the BDT score value to be above 0.68 (0.3) in the τ_{lep}-τ_{lep} (τ_{lep}-τ_{had}) channel, resulting in 49% (51%) efficiency with respect to the full VBF region for the signal and 3.6% (2.1%) for the background processes. It was checked that the BDT score does not affect the mean value of the Optimal Observable

and the modelling of the $\mathcal{OO}$ for the background processes is in agreement with data in various control regions.

The systematic uncertainties under consideration follow those of the previously described $H \to \tau\tau$ analysis with the addition of a contribution from the re-weighting of the signal sample to the non-SM $\tilde{d}$ values. The most important systematic contributions were however found to arise from energy scales of jets, hadronic τ decays and electrons, while the analysis is statistically limited.

A maximum-likelihood fit to the Optimal Observable distribution is performed to extract the value of $\tilde{d}$. Information from the control regions is included to constrain background normalisations and nuisance parameters. The normalisation of the Higgs boson signal is allowed to float in the fit. A set of signal templates with varying values of $\tilde{d}$ has been created by re-weighting the nominal signal sample and the likelihood value is computed for each $\tilde{d}$ hypothesis using the corresponding signal template while keeping the backgrounds unchanged. Nuisance parameters are also profiled to their best-fit values, including those describing systematic uncertainties.

Results and summary

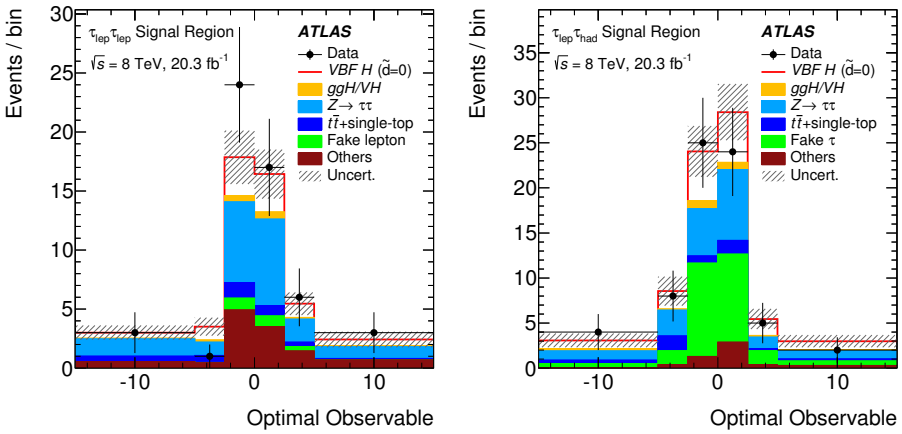
After performing the aforementioned template fits, each of the $\tilde{d}$ values under consideration has a negative log-likelihood (NLL) value assigned. The best estimate for $\tilde{d}$ is the NLL minimum and the approximate central confidence interval at 68% confidence level is defined for $\Delta NLL = NLL_{\text{min}} - 0.5$. The obtained limits of the 68% confidence interval for the $\tilde{d}$ values are $[-0.11, 0.05]$ ($\tilde{d}$ values below and above this interval are excluded at 68% confidence level), fully consistent with no CP-violation in this sector, as predicted by the Standard Model. The obtained 68% confidence-level interval is a factor 10 better than that of the previous ATLAS analysis [176]. Figure 5.6 shows the post-fit distributions of the Optimal Observable in both considered analysis channels, for the $\tilde{d} = 0$ hypothesis. The obtained signal strength is $\mu = 1.55^{+0.87}_{-0.76}$, also fully consistent with the SM.

5.2.3 Measurement of $H \to \tau\tau$ cross-section

The cross-section for the Higgs boson production is measured in the $H \to \tau\tau$ decay channel [164] with the early-Run-2 dataset collected in 2015 and 2016 with $\sqrt{s} = 13$ TeV and corresponding to the integrated luminosity of 36 fb^{-1}. Many ideas from the Run-1 $H \to \tau\tau$ evidence analysis (Sec. 5.2.1 and Ref. [37]) are also used here. The di-τ invariant mass, $m_{\tau\tau}^{\text{MMC}}$, computed by the MMC is used as the discriminating variable in this measurement.

Samples for analysis

A sample of collision-data events was selected using a set of triggers similar to that of Run-1 study. Single-lepton triggers with p_{T} thresholds between 21 and 27 GeV were used as well as di-lepton ones with thresholds for the leading (sub-leading) lepton varying between 15 and 24 (10 and 18) GeV. In addition, di-τ triggers with p_{T} thresholds of 40 and 30 GeV were also in place. The p_{T} thresholds were increased in the 2016 data taking due to higher luminosity and larger number of pile-up interactions per bunch crossing, with the average of 25, compared to the 2015 value of 14.



(a) (b)

Figure 5.6: Distributions of the Optimal Observable in the signal region for the τ_{lep}-τ_{lep} (left) and τ_{lep}-τ_{had} channel (right), after performing the global fit for the $\tilde{d} = 0$ hypothesis [163].

Samples of simulated events were produced for the examined signal and the expected background contributions. Signal processes of interest included four Higgs production channels, ggF, VBF, VH and top-associated $t\bar{t}H$. All but the last one were produced with the Powheg-Box [177–179] generator whereas the last one was produced with MadGraph5_aMC@NLO. The ggF production was simulated at next-to-leading order (NLO) and re-weighted to next-to-next-to-leading order (NNLO) in QCD in Higgs rapidity. The VBF and VH were simulated at NLO accuracy, similarly to the $t\bar{t}H$. The Pythia 8 generator was used to simulate parton showering, hadronisation and underlying event in these processes. Background samples of electroweak boson decays (accompanied by jets) were produced with the Sherpa 2.2.1 [106] generator and the background involving top-quark decays was generated with Powheg-Box interfaced with Pythia 6. Contrary to the Run-1 analysis, no $Z/\gamma^* \to \tau\tau$ hybrid (embedded) sample was used in this study.

Event selection and analysis categories

Similarly to the Run-1 studies, three analysis channels based on τ-pair decays are singled out, namely the τ_{lep}-τ_{lep}, τ_{lep}-τ_{had} and τ_{had}-τ_{had}. In each of these channels exact numbers of *loose* light leptons or τ's are required, with leptons fulfilling isolation criteria. *Medium* τ identification requirement together with p_{T} threshold of 30 GeV is imposed in the τ_{lep}-τ_{had} channel whereas in the τ_{had}-τ_{had} channel two *tight* τ's ere expected. Opposite charges of the two decay legs are required together with $E_{\text{T}}^{\text{miss}}$ above 20 GeV (raised to 55 GeV in the same-flavour τ_{lep}-τ_{lep} channel with the addition of $E_{\text{T}}^{\text{miss HPTO}} > 55$ GeV requirement to suppress $Z \to \ell\ell$ background). In the τ_{lep}-τ_{lep} channel the di-τ invariant mass estimated with the collinear approximation should not be smaller than 66 GeV and the di-lepton mass should lie within $30 < m_{\ell\ell} < 75$ (100) GeV in the same-flavour (different-flavour) channel. The angular separation, ΔR of the τ's in all channels should be below 2.5 (2.0 in the τ_{lep}-τ_{lep} channel) and above 0.8 in the τ_{had}-τ_{had} channel. Pseudorapidity distance

between the τ's should be below 1.5 to suppress non-resonant backgrounds. The visible momenta fractions (from collinear approximation) must not be lower than 0.1 and can not exceed 1.4 (1.0 in the τ_{lep}-τ_{lep} channel) to reject processes with $E_{\text{T}}^{\text{miss}}$ not compatible with the di-τ decay. In each of the channels an accompanying jet with $p_{\text{T}}^{j_1} > 40$ GeV (70 GeV in τ_{had}-τ_{had} channel) is required to select VBF processes and to suppress $Z/\gamma^* \to \tau\tau$ background in selection of ggF production. b jets are vetoed in the τ_{lep}-τ_{lep} and τ_{lep}-τ_{had} channels in addition.

Two signal-sensitive topologies are used by defining two separate analysis categories, the *VBF* and the *boosted* one, as in Run-1 analysis. The VBF category is characterised by another high-p_{T} jet ($p_{\text{T}}^{j_2} > 30$ GeV) in the opposite detector hemisphere with respect to the first jet ($|\Delta\eta_{jj}| > 3$) and with invariant mass of the di-jet system above 400 GeV. The τ's and leptons are required to lie between these two jets. This category contains VBF events with a 30% admixture of ggF production. The boosted category aims for exploiting the topology of ggF events with additional recoiling jets by requiring the p_{T} of the $\tau\tau$ system to be above 100 GeV. This requirement can strongly reduce the $Z/\gamma^* \to \tau\tau$ background contribution. In addition, the VBF-category criteria need to be failed for the event to be selected in the boosted category. In addition to the ggF events it contains $10 - 20\%$ of VBF and VH admixture. These categories are then split into 13 exclusive signal regions based on the values of $\tau\tau$-system p_{T}, angular separation of τ pairs and the di-jet invariant mass, to improve the analysis sensitivity.

Background estimation

Estimation of background for the multiple analysis categories and channels requires separate strategies for each of them.

Due to technical limitations, the embedded hybrid $Z/\gamma^* \to \tau\tau$ sample was not available by the time of performing this analysis. A regular MC sample is then used for estimation of this background. The modelling of this contribution is validated using dedicated validation regions containing $Z \to \ell\ell$ events with kinematics similar to that of $Z/\gamma^* \to \tau\tau$ in the respective signal regions. This is achieved by following the τ_{lep}-τ_{lep} same-flavour selection but dropping the $E_{\text{T}}^{\text{miss}}$ and collinear-approximation di-τ mass requirements and inverting the di-lepton mass cut to $m_{\ell\ell} > 80$ GeV. Distributions of several variables correlated with the $m_{\tau\tau}^{\text{MMC}}$ are checked and a general agreement of their description in the SHERPA Monte Carlo is found.

The backgrounds from light and heavy-flavour jets mimicking light leptons or τ's are estimated using data-driven methods. For the former (important in the τ_{lep}-τ_{lep} channel), inverted sub-leading lepton isolation criteria are applied to construct the control regions and top and $Z \to \ell\ell$ backgrounds are subtracted using simulations. Correction factors scaling these contributions are obtained from same-sign control regions. In case of jets mimicking hadronic τ decays in the τ_{lep}-τ_{had} channel the fake-factor method is used to obtain their contribution, as in Run-1 analysis. The multi-jet background estimation in the τ_{had}-τ_{had} channel exploits a control region with two-track τ candidates failing opposite-sign requirement. Template extracted from that region is then re-weighted based on the azimuthal angular separation between the τ candidates by comparing the templates from this region to those from the OS region but with one of the τ's failing the *tight* identification requirements but fulfilling the *medium* ones.

Backgrounds with genuine hadronic and leptonic τ decays as well as with prompt leptons are estimated from simulations and, in case their contribution is significant, their

normalisation is constrained by the observed event yields in control regions, such as the top control-region, with inverted b-jet veto and $Z \to \ell\ell$ control regions with di-lepton mass required to lie within the 80 - 100 GeV mass window.

Systematic uncertainties

The shape of $m_{\tau\tau}^{\text{MMC}}$ distribution as well as the expected signal and background yields are affected by systematic uncertainties. The most pronounced uncertainty source in this analysis is the theoretical modelling of signal production process that affects the analysis acceptance in addition to the uncertainty related to the choice of parton-shower model. Another important uncertainty sources are related to the energy-scale calibration of jets and $E_{\text{T}}^{\text{miss}}$, estimation of background and finally the τ energy scale and identification efficiencies.

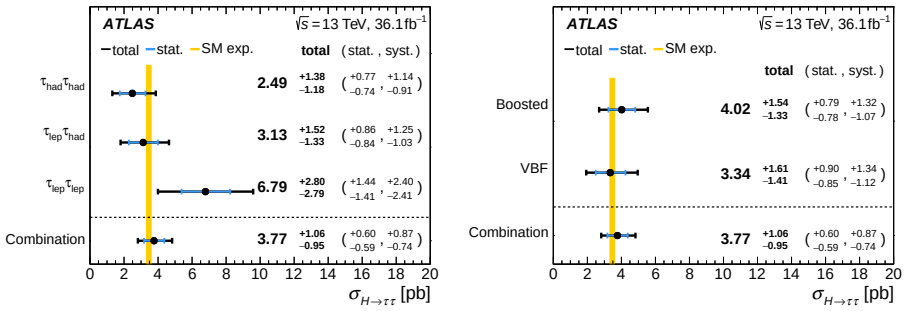
Results and summary

A maximum-likelihood fit is performed on the $m_{\tau\tau}^{\text{MMC}}$ distributions to extract the $H \to \tau\tau$ cross-section. All the signal and control regions are used in the fit and systematic uncertainties are parametrised as nuisance parameters. The observed (expected) signal significance in this measurement is 4.4 (4.1) standard deviations, compatible with Higgs boson at $m_H = 125$ GeV. The combination with Run-1 results gives 6.4 (5.4) standard-deviation observed (expected) significance.

The inclusive $\sigma_{H\to\tau\tau} = \sigma_H \times \mathcal{B}(H \to \tau\tau)$ parameter is extracted from the fit. The production cross-section contains all the relevant production processes (ggF, VBF, VH and $t\bar{t}H$) with relative contributions assumed to be as predicted by the SM. The measured value is $\sigma_{H\to\tau\tau} = 3.77^{+0.60}_{-0.59}(\text{stat})^{+0.87}_{-0.74}(\text{syst})$ pb, in agreement with the SM prediction of $\sigma^{\text{SM}}_{H\to\tau\tau} = 3.36 \pm 0.13$ pb [32]. The signal strength is $\mu_{H\to\tau\tau} = 1.09^{+0.18}_{-0.17}(\text{stat})^{+0.31}_{-0.25}(\text{syst})$. In addition, measurements in the two dominant production modes and in the three individual analysis channels were made. The results are summarised in Fig. 5.7, showing a good agreement with the SM predictions for the individual production channels as well as for the combined result. The post-fit distributions of the $m_{\tau\tau}^{\text{MMC}}$ distributions compared with data for the individual analysis categories are presented in Fig. 5.8.

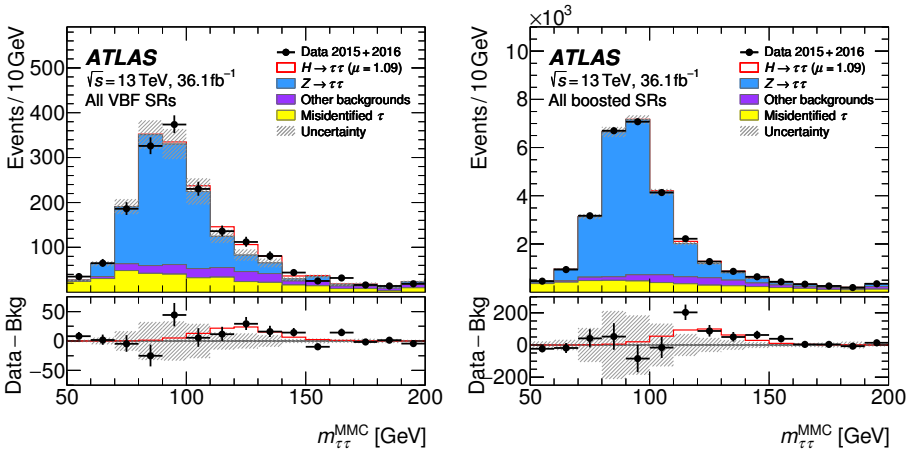
5.2.4 Summary

The presented results of all three analyses in the $H \to \tau\tau$ decay channel are in agreement with the Standard Model expectations. The CMS collaboration has performed similar measurements. In the Run-1 search for the $H \to \tau\tau$ process a signal excess with local significance of 3.2 standard deviations was found and the corresponding signal strength was 0.78 ± 0.27 times the SM expectations [38]. A direct observation of the $H \to \tau\tau$ decay was performed by the CMS experiment with 5.9 standard-deviations significance using combined Run-1 and Run-2 data [180]. No cross-section is quoted in that publication, however the observed signal strength is 1.09 ± 0.27, in agreement with the ATLAS Run-2 result. The observed signal strength values of both experiments are summarised in Table 5.3.



(a) (b)

Figure 5.7: Measured $\sigma_{H\to\tau\tau}$ values in the individual analysis channels (left) and in the individual production modes (right). Combined fit results are also shown and the SM expectations are overlaid. The total $\pm1\sigma$ uncertainty is shown as black error bars whereas the statistical uncertainty is indicated by blue bars [164].



(a) (b)

Figure 5.8: Distributions of the di-τ invariant mass, $m_{\tau\tau}^{\mathrm{MMC}}$ after global fit in the VBF (left) and boosted (right) categories summed over all signal regions. Bottom panels show the difference between the data events and expected background. The observed Higgs boson signal at $\mu_{H\to\tau\tau} = 1.09$ is shown with the solid red line. Total uncertainty is marked with hatched bands [164].

	ATLAS $\mu_{H\to\tau\tau}$	CMS $\mu_{H\to\tau\tau}$
Run 1	$1.43^{+0.43}_{-0.36}$	0.78 ± 0.27
Run 2	$1.09^{+0.36}_{-0.30}$	$1.09^{+0.27}_{-0.26}$

Table 5.3: Observed signal-strength parameter values for the Run-1 and Run-2 searches for the $H \to \tau\tau$ decays in the ATLAS and CMS experiments.

5.3 Observation of $H \to b\bar{b}$ and VH production

The most frequent decay mode of the Higgs boson is the $H \to b\bar{b}$ with the expected branching fraction of about 58% [181]. This measurement [79] aims at probing the dominant Higgs boson decay mode that can constrain the overall Higgs boson decay width [32, 182]. Because of the overwhelming multi-jet backgrounds at the LHC, the dominant production modes, ggF and VBF, are difficult to access experimentally in this decay channel. Instead, the VH production mode can be used with the additional associated vector boson which provides an additional clean experimental signature with its leptonic decays.

Previously, searches for $H \to b\bar{b}$ decays in this production mode were performed at Tevatron by the CDF and D0 experiments and showed an excess of events at 2.8σ significance level for Higgs boson of 125 GeV mass [183]. Evidence was found by the ATLAS and CMS experiments after combining the results of this search in Run-1 data [184, 185] with those obtained with the 2015-2016 dataset [186, 187] with the observed (expected) significances of 3.6 (4.0) and 3.8 (3.8) standard deviations, respectively. Searches were also performed in the VBF [188, 189] and $t\bar{t}H$ [190, 191] production modes but with much lower sensitivity.

5.3.1 Data and simulation samples

The described analysis uses the subset of Run-2 ATLAS data collected in 2015-2017 period corresponding to 79.8 fb^{-1}, collected at $\sqrt{s} = 13$ TeV with an average number of inelastic pp collisions in each bunch crossing totalling to 32. Events are selected with a set of triggers aiming to exploit the presence of the vector boson. They include $E_{\mathrm{T}}^{\mathrm{miss}}$-based selection with thresholds varying from 70 to 110 GeV depending on the pile-up conditions to select events with $ZH \to \nu\nu b\bar{b}$ decays as well as $WH \to \mu\nu b\bar{b}$ processes[1]. Single-electron triggers with thresholds of 24-26 GeV are used for the $WZ \to e\nu b\bar{b}$ and $ZH \to eeb\bar{b}$ decays whereas single-muon triggers with p_{T} thresholds of 20 to 26 GeV aim to select $ZH \to \mu\mu b\bar{b}$ signatures.

Signal Monte Carlo samples were produced with the Powheg-Box program interfaced with Pythia 8 for parton shower and hadronisation. This was also the case for the background samples of top-quark decays. Samples of vector-boson (including di-boson) decays with accompanying jets were produced with the Sherpa 2.2.1 event generator. The EvtGen [192] program was used to simulate the decays of b and c hadrons for all samples except the Sherpa ones.

5.3.2 Selection and categorisation of events

The selection criteria in this analysis are constructed to efficiently exploit the various signal event topologies. Three selection categories are defined, based on the number of final-state light charged leptons: 0-lepton, 1-lepton and 2-lepton, 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. Common requirement to all categories is that two b-tagged jets are present in the event with at least one with $p_{\mathrm{T}} > 45$ GeV. In addition, as the signal-to-background ratio is expected to increase for large Higgs-boson transverse momenta [193, 194], this analysis selects events with high p_{T} of the vector boson, $p_{\mathrm{T}}^{V} > 150$ GeV. In the 2-lepton category also the medium p_{T}^{V} region

[1] The $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers don't include muons in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, so in the $W \to \mu\nu$ case they effectively select high W-p_{T} events.

with $75 < p_{\mathrm{T}}^{V} < 150$ GeV is used. Additional untagged jets are allowed by the signal selection and the events are further split into 2-jet and 3-jet categories. For the 0 and 1-lepton category only up to 1 additional jet is allowed whereas in the 2-lepton 3-jet category events with any number of additional jets are accepted. Eight signal regions emerge from this selection procedure.

In the 0-lepton category, in addition to the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger selection, the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be above 150 GeV. No *loose* leptons are allowed. The scalar sum of jet transverse momenta, H_{T} is required to exceed 120 (150) GeV in the 2-jet (3-jet) category. Angular separation between the direction of $E_{\mathrm{T}}^{\mathrm{miss}}$ and di-jet system should be above 120° and the angle between the $E_{\mathrm{T}}^{\mathrm{miss}}$ and track-based missing transverse momentum should be below 90°. Azimuthal angle separation between the two b jets should also not exceed 140°.

Events containing exactly one *tight* electron (muon) with p_{T} above 27 (25) GeV are accepted in the 1-lepton category after passing a single-electron ($E_{\mathrm{T}}^{\mathrm{miss}}$) trigger selection criteria for the electron (muon) subcategory. Events with additional *loose* leptons are rejected. A requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV is imposed in the electron subcategory to reduce background from multi-jet production. Events are further categorised into a signal region or a control region enriched in W+heavy-flavour (W+HF) events with the invariant mass of the two b jets, m_{bb}, and with the invariant mass constructed from the momenta of two b jets, lepton and neutrino (approximated from $E_{\mathrm{T}}^{\mathrm{miss}}$ and by constraining the $\ell\nu$ system to the mass of W boson), denoted m_{top} and approximating the top quark invariant mass. They are required to be $m_{bb} > 75$ GeV ($m_{bb} < 75$ GeV) and $m_{\mathrm{top}} < 225$ GeV ($m_{\mathrm{top}} > 225$ GeV) for the signal (control) region.

In the 2-lepton category the events are first required to be selected by the single-electron or single-muon triggers. Afterwards, events with two *loose* leptons, out of which at least one should have $p_{\mathrm{T}} > 27$ GeV are accepted. In addition, the invariant mass of the di-lepton system should be compatible with Z-boson mass. Events with leptons of the same flavour enter the signal region whereas different-flavour ($e\mu$) events are selected for the top-quark control region.

5.3.3 Multivariate analysis

In the previous round of this analysis [186] the construction of the multivariate discriminant based on the BDT algorithms was established and is used in the described measurement. The BDTs are trained in each of the signal categories separately. Variables used across all categories include the invariant mass of the b jets, vector-boson p_{T}^{V} (equivalent to $E_{\mathrm{T}}^{\mathrm{miss}}$ in 0-lepton category), $E_{\mathrm{T}}^{\mathrm{miss}}$, individual p_{T} of b jets and their angular separation as well as the azimuthal-angle separation between the vector boson and di-b-jet system. To enhance the sensitivity in the 3-jet category the p_{T} of the third jet and invariant mass of the di-b-jet+jet system are added to the training. Additional suppression of the $t\bar{t}$ background in the 1-lepton category is achieved by including the information on reconstructed top-quark mass (assuming the event is a $t\bar{t}$ decay) and the azimuthal angle between the lepton and closest b jet. The output from the BDT algorithm is used as the final discriminating variable in search for the $H \to b\bar{b}$ signal.

5.3.4 Estimation of background

The background from multi-jet production is estimated with data-driven methods. In the 0- and 2-lepton categories it is extracted from fits to data with a functional-form

description of its shape and other background processes estimated with simulations. In the 0-lepton category the spectrum of angular separation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and closest jet is used whereas in the 2-lepton category di-lepton invariant mass with same-sign selection is used in the fit. In both these cases the multi-jet contribution was found to be negligible. In the 1-lepton category the normalisation of this background is estimated with fits to W-boson transverse mass distributions in all signal categories separately, with relaxed-lepton identification criteria. The shape of the BDT distribution is also obtained from the mentioned control regions.

The estimation of other background sources in the final selected signal regions is made using simulated samples with normalisations derived in control regions and extrapolated to the relevant signal regions. The shapes and normalisations of the Z+heavy-flavour jets (Z+HF) can be constrained in the 2-lepton category, whereas for the W+HF processes there is a dedicated control region constructed in the 1-lepton category. The background resulting from top-quark decays in the 0- and 1-lepton category is estimated separately from the 2-lepton category due to different regions of phase-space being probed. In the 0- and 1-lepton categories the constraints come from the W+HF control regions whereas in the 2-lepton category dedicated top ($e\mu$) control regions were defined.

5.3.5 Systematic uncertainties

Systematic uncertainties of this measurement affect the shapes and normalisations of the BDT distributions of both the signal and backgrounds. They are dominated by the data-to-simulation scale factors related to b-jet tagging. Background modelling uncertainty also constitutes an important systematic contribution as well as the modelling of signal acceptance.

5.3.6 Statistical analysis and results

The signal strength parameter μ_{VH}^{bb} is extracted using a simultaneous maximum-likelihood fit in all the signal and control regions. Systematic uncertainties are parametrised as nuisance parameters. This is also the case for the floating background normalisations of the largest backgrounds from $t\bar{t}$ and W/Z+HF events.

After combining all analysis categories the statistical significance of the signal excess over the SM background is 4.9 standard deviations (4.3 expected) and the resulting signal strength is $\mu_{VH}^{bb} = 1.16 \pm 0.16(\mathrm{stat})^{+0.21}_{-0.19}(\mathrm{syst})$. The post-fit distributions of the BDT output score in the six high-p_{T}^{V} analysis regions are shown in Fig. 5.9.

The multivariate analysis was cross-checked with a parallel study based on the m_{bb} variable instead. This check returned a compatible result with significantly lower sensitivity, as expected.

5.3.7 Combinations with other searches

The statistical significance of the result of the ATLAS Run-2 VH, $H \to b\bar{b}$ search alone was not sufficient to declare the 5σ observation. Therefore, a combination of results with searches for $H \to b\bar{b}$ decays in other production modes and with Run-1 searches was made, resulting in the observed signal excess significance of 5.5 standard deviations (5.4 expected) with signal strength of $\mu_{H\to b\bar{b}} = 1.01 \pm 0.12(\mathrm{stat})^{+0.16}_{-0.15}(\mathrm{syst})$, resulting in the direct observation of the $H \to b\bar{b}$ decay channel. The direct observation of VH production

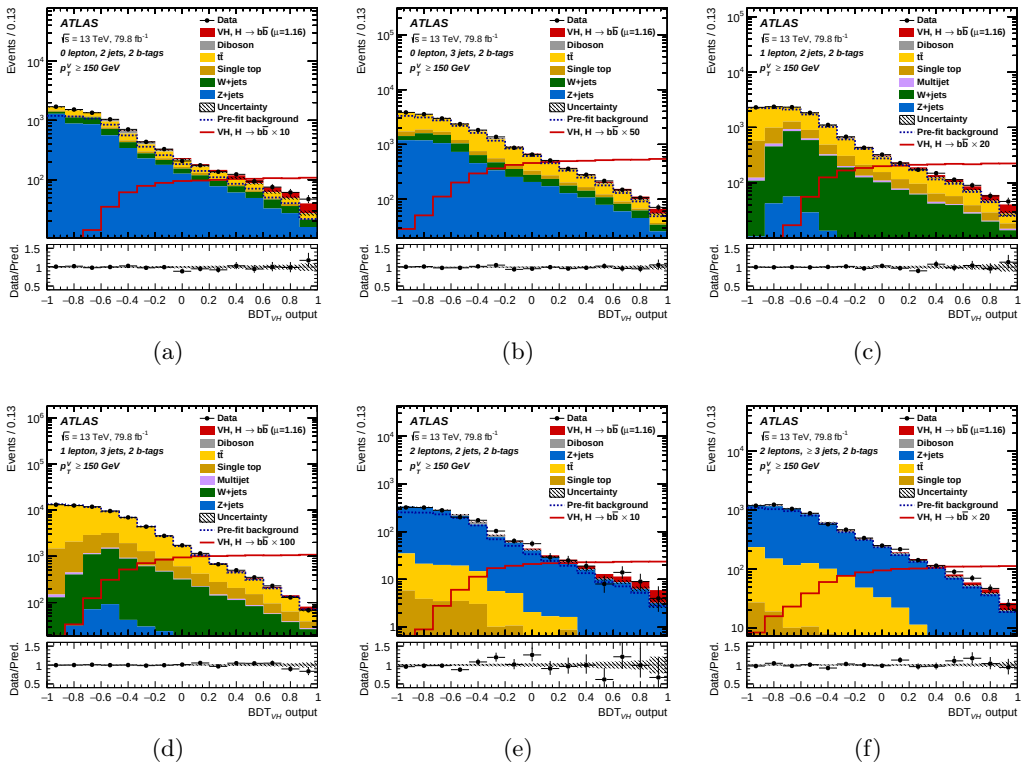


Figure 5.9: BDT distributions in 2-jet (top row) and 3-jet (bottom row) categories in the 0-lepton (left column), 1-lepton (middle column) and 2-lepton (right column) categories. The Higgs boson signal with mass of 125 GeV is also shown with signal strength of 1.16 (filled area) and 100 (solid line) times the SM expectations. All normalisations are obtained with the final fit [79].

mode was completed by combining the searches in several Higgs decay channels, including $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$, resulting in the observed excess significance of 5.3 standard deviations (4.8 expected) and signal strength of $\mu_{VH} = 1.13 \pm 0.15(\text{stat})^{+0.18}_{-0.17}(\text{syst})$.

5.3.8 Summary

The described VH, $H \to b\bar{b}$ analysis had a crucial impact on the direct observation of the most frequent Higgs-boson decay mode into $b\bar{b}$ pairs in combinations of the ATLAS results. In addition, the direct observation of the vector-associated Higgs boson production mode was provided. All the presented results are consistent with the Standard-Model predictions for the Higgs sector.

A very similar set of results was meanwhile obtained by the CMS experiment [80] yielding a 4.8 standard-deviations significance for the VH search combinations and 5.6 standard deviations excess in the combination of $H \to b\bar{b}$ decay searches. The comparison of the signal-strength parameters obtained by ATLAS and CMS is presented in Table 5.4.

	$\mu_{H\to b\bar{b}}$	μ_{VH}
ATLAS	$1.01^{+0.20}_{-0.19}$	1.13 ± 0.23
CMS	1.04 ± 0.20	1.01 ± 0.22

Table 5.4: Observed signal-strength parameter values for the Run-1 and Run-2 searches for the $H \to b\bar{b}$ decays and VH production in the ATLAS and CMS experiments.

5.4 Summary

The measurements related to the Standard Model processes are being performed in the LHC experiments showing no significant deviations from theory predictions so far. Their precision is increasing and so is the understanding of the Standard Model physics. The scalar sector of the SM, a relatively new field of experimental particle physics, is under constant exploration. The results of measurements, including the presented ones, show that our understanding of this sector is correct, and that, in particular, the Yukawa mechanism for generating fermion masses can be confirmed. Further measurements with increased statistics of data will soon bring even more insight into that sector, thanks to more precise studies of the processes described in this Chapter, but also thanks to the searches for $H \to \mu\mu$ and $H \to c\bar{c}$ that are conducted in the LHC experiments [195–197].

Chapter 6

Searches for Beyond-Standard-Model Higgs bosons with third-generation-fermion final states

In addition to precise measurements of various Standard Model properties, the principal objective of the LHC has always been searching for the unknown. In the current state of knowledge it is apparent that the SM is not sufficient to describe all the observed phenomena in the world of particle physics, including the nature of Dark Matter, mater - anti-matter asymmetry or neutrino masses to name but a few, and does not attempt to describe quantum gravitation. The LHC experiments may be able to provide some hints on the Beyond-Standard-Model physics by means of discovering previously unobserved particles or measuring anomalous couplings that may, for example, significantly change cross-sections of measured processes with respect to their SM expectations.

A simple extension of the SM is the Two-Higgs-Doublet Model introducing an additional doublet of scalar fields to the Higgs sector. Its existence would naturally emerge from supersymmetric theories, however other motivations for it include potential new sources of the CP violation, needed to explain the mater-antimatter asymmetry [17]. This Chapter focuses on the searches in the BSM Higgs sector.

6.1 Search for heavy neutral Higgs boson produced in association with b quarks and decaying into b-quark pair

The Higgs boson couplings to the b quarks are strong already in the Standard Model, but a significant enhancement is expected for the BSM scenarios, as stated in Sections 2.1.4 and 2.2.4. As can be seen from Table 2.2, the $b\bar{b}$ is the most sensitive BSM Higgs decay channel to the type-Y (*flipped*) 2HDM. In addition, it can also be used to constrain type-II searches, yet the sensitivity in the $\tau\tau$ channel is expected to be better (see Sec. 6.2).

A search for the heavy neutral Higgs boson, H/A, produced with at least one associated b quark and decaying into $b\bar{b}$ pair performed by the ATLAS experiment [53] is described here. This search has no sensitivity for distinction between the CP-even and CP-odd H and A states as their decay kinematics are nearly identical. Also, no prior assumption on their mass degeneracy is made. The b-associated production is considered because of the cleaner experimental signature it provides, as the multi-jet backgrounds are still overwhelming even at masses well above the SM Higgs mass. The production process is illustrated in Fig. 2.8.

6.1.1 Data and simulation event samples

This search uses the dataset corresponding to 27.8 fb^{-1} collected in 2015 and 2016[1] at the pp collision energy of $\sqrt{s} = 13$ TeV. Triggers selecting b-tagged jets were employed, including single b-jet trigger with p_{T} threshold of 225 GeV and di-b-jet trigger accepting two b-jets with transverse momenta over 150 and 50 GeV. The online b-tagging operating points of 79% (60%) efficiency were used for the single-b-jet trigger in 2015 (2016) whereas the 2015 (2016) di-b-jet selection used 72% (60%)-efficiency working points.

Expected signal events were simulated with the Sherpa 2.2.0 generator for hypothetical H/A masses between 450 and 1400 GeV with widths typical to 2HDM-flipped models, no larger than $10-15\%$ of the experimental $b\bar{b}$ mass resolution. The backgrounds are estimated using data-driven technique, however, to help developing analysis strategy, a multi-b-jet MC sample was also produced with the Sherpa 2.1.1 program. Additionally, Powheg+Pythia 6 and Madgraph [198] were used to simulate $t\bar{t}$ and $Z \to b\bar{b}$+jets samples, respectively.

6.1.2 Selection of events

Following the trigger selection, the events are accepted for the actual analysis if they contain at least two jets tagged as b-jets at the 70%-efficiency working point with transverse momenta above 160 and 60 GeV for the leading and sub-leading jet, respectively. In addition, at least one additional jet with $p_{\mathrm{T}} > 25$ GeV is required, with angular separation from the closest of the two leading b-jets, $\Delta R_3^{\mathrm{min}} > 0.8$. Events are further selected into signal region, called *bbb*, if the third jet fulfils the same b-tagging criteria as the two leading ones. Otherwise the event enters background region, *bbanti*. The events in both regions are further classified based on the total jet multiplicity into categories of 3-jets, 4-jets and 5-jets with the last one containing events with five jets or more. This classification improves the analysis sensitivity as the events with larger number of jets more likely contain final-state

[1] Because of a technical problem with b-jet triggers a part of data was not recorded.

radiation (FSR) jets that distort the spectrum of the invariant mass of the two leading b-jets as well as their transverse momenta.

6.1.3 Invariant mass rotation

The natural candidate for the variable for discriminating the signal against background would be the invariant mass of the potential Higgs decay products - the two b-jets in this case, m_{bb}. However, at higher $m_{H/A}$ the high-p_{T} jets from the H/A decay produce additional FSR and the m_{bb} spectrum is smeared out and becomes more difficult to be distinguished from the background. On the other hand, there are events with small amounts of FSR and therefore with the value of m_{bb} closer to the true value of the hypothetical heavy Higgs-boson mass. Selection of such events is helped with a rotation of the tensor built of three-dimensional distributions of the m_{bb} and transverse momenta of the two leading b-jets, $p_{\mathrm{T}1}$, $p_{\mathrm{T}2}$ in the signal samples, based on the *principal component analysis* method [199, 200]. Upon diagonalisation the first principal axis of the tensor defines the m'_{bb} variable, use of which increases the sensitivity of the described search. In addition, the second and third principal axes define the $p'_{\mathrm{T}1}$ and $p'_{\mathrm{T}2}$ variables, respectively. The *new* variables are linear combinations of the *old* ones, with coefficients depending on the considered mass point. The dominant component of the m'_{bb} is the m_{bb}, regardless of the mass point. The final selection criteria involve requirements imposed on these new variables, namely $p'_{\mathrm{T}1} > -10$ GeV and $p'_{\mathrm{T}2} > -50$ GeV. This transformation is performed separately for each considered Higgs mass point and therefore the final discriminating variable as well as selection criteria are mass-point dependent. An illustration of the effect of the described transformation can be observed in Fig. 6.1 presenting two-dimensional distributions of the m_{bb} and $p_{\mathrm{T}1}$ variables before and after transformation with $p'_{\mathrm{T}1}$ selection shown with dashed line. Similar behaviour is observed in the $p_{\mathrm{T}2}$ vs m_{bb} distributions and in other mass points.

6.1.4 Statistical analysis

The search for the signal is performed using binned maximum-likelihood fit to the data using the m'_{bb} variable. The fit is performed separately in each of the considered mass points and simultaneously in all the analysis regions and categories. The shapes and normalisations in each category are sums of signal and background contributions. The shapes of signal distributions are obtained from simulations and their normalisations constitute the parameters of interest of the fit. The normalisation of signal histograms entering the fit are normalised to event yields expected for a 1 pb signal. The shapes and normalisations of background are estimated with the fit, with the constraint that the background shapes in the *bbb* and *bbanti* regions are identical modulo a second-order polynomial correction factor. The backgrounds are estimated separately in each of the jet-multiplicity categories.

The fit model includes nuisance parameters that account for systematic variations (see Sec. 6.1.5) of the shapes and normalisations of the signal template histograms as well as of the limited statistics of the MC samples.

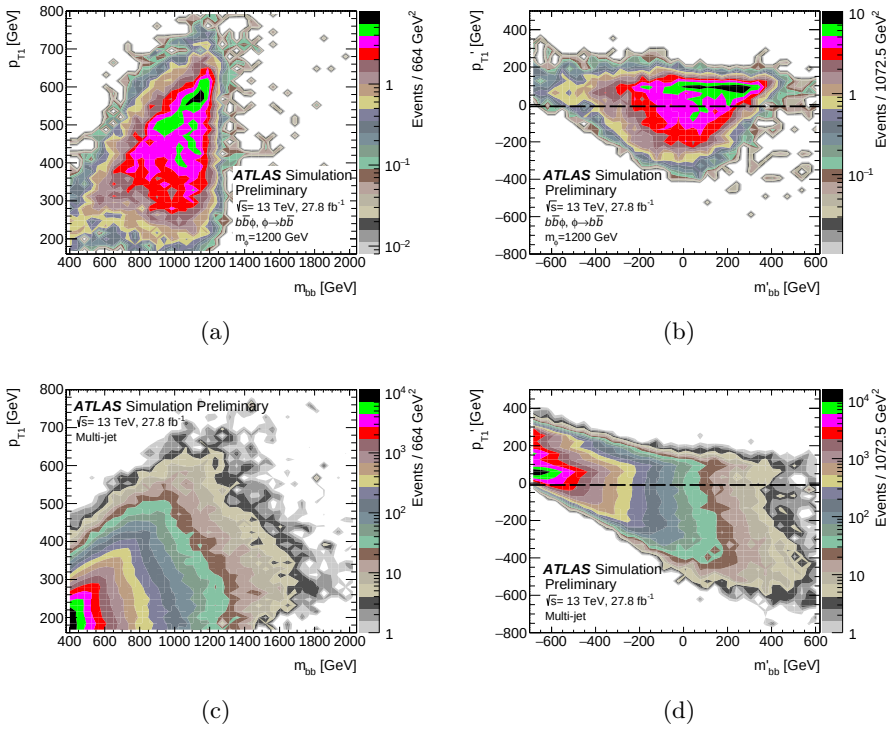


Figure 6.1: Two-dimensional distributions of the p_{T1} vs m_{bb} before (left) and after invariant-mass rotation (right) before signal/background region categorisation, summed over jet-multiplicity categories for the $m_{H/A} = 1200$ GeV mass point. The top row presents the expected signal whereas the bottom row shows the distributions of multi-jet simulated events. The signal distributions assume signal cross-section of 1 pb and the normalisation of multi-jet MC is arbitrary. Distributions of p_{T2} vs m_{bb} and in other Higgs mass points show similar behaviour [53].

6.1.5 Systematic uncertainties

The shapes and normalisations of the predicted m'_{bb} distributions are subject to systematic uncertainties. Regarding signal, the dominant experimental systematic contribution arises from the b-tagging efficiency correction factors both at trigger and offline-selection level. The uncertainties of the online b-tagging correction factors increase with jet p_{T} and, consequently, with the considered Higgs boson mass point. Jet energy scale and resolution uncertainties are also considerable in this analysis. In addition, signal modelling is affected by the choice of generator for the hard process and by the choice of showering and hadronisation models. This contribution was estimated by comparing the default signal samples with auxiliary ones produced with MadGraph5_aMC@NLO interfaced with Pythia 8.

The background estimation is entirely data-driven, with the assumption on the ratio of background shapes in the *bbb* and *bbanti* regions being a second-order polynomial at the most. This was extensively tested including also the *F-test* method [201], showing that, actually, first-order polynomial is sufficient for mass points below 1200 GeV in the 3-jet category and below 800 GeV in the 4- and 5-jet categories, whereas second-order polynomial is sufficient elsewhere.

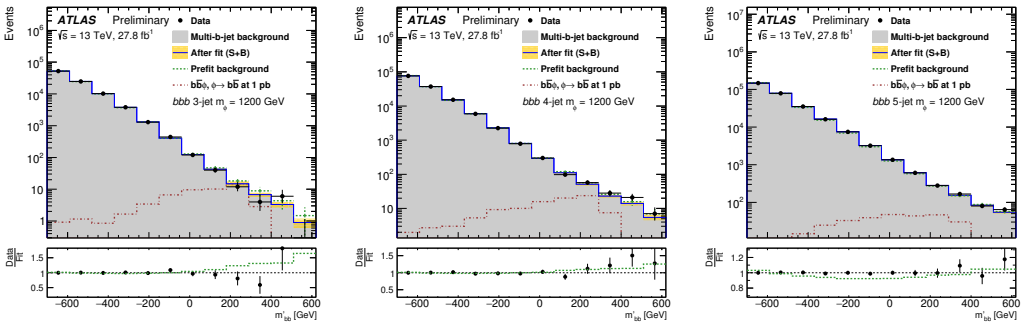
6.1.6 Results and interpretation

The results of the search for heavy neutral Higgs boson, produced in association with b quarks and decaying into $b\bar{b}$ pair show no significant excess of the signal over background for any of the considered mass points. Example values of systematic uncertainty contributions as estimated by the fit are presented in Table 6.1 for mass points of $m_{H/A} = 600$ and 1200 GeV. The dominant total uncertainty contribution in all mass points is the statistical uncertainty.

Source of uncertainty	$m_{H/A} = 600$ GeV $\Delta(\sigma \times B)$ [pb]	$m_{H/A} = 1200$ GeV $\Delta(\sigma \times B)$ [pb]
Total	0.8	0.29
Statistical	0.77	0.26
Systematic	0.2	0.11
Experimental uncertainties		
Jet-related	0.05	0.05
B-tagging (offline)	0.12	0.05
B-trigger	0.04	0.05
Luminosity	0.02	0.01
Theoretical and modelling uncertainties		
Generator	0.03	0.03
PDF	0.08	0.04
MC statistical	0.09	0.04

Table 6.1: Grouped systematic contributions to the uncertainty on best-fit value of $\sigma \times \mathcal{B}$.

Post-fit plots illustrating the results for the $m_{H/A} = 1200$ GeV mass point for each of the jet-multiplicity categories in the bbb region are shown in Fig. 6.2. With no significant



(a) (b) (c)

Figure 6.2: Distributions of m'_{bb} in the bbb region after final fit for the $m_{H/A} = 1200$ GeV mass point in 3-jet (left), 4-jet (middle) and 5-jet (right) categories. Pre-fit background is also shown (green dashed line) together with the expected signal for 1 pb cross-section (dashed red line) [53].

signal excess found, upper limits on the production of the heavy Higgs boson in association with b quarks and decaying into b-quark pairs are set. The observed and expected limits for the cross-section of this process at the 95% confidence level are presented in Fig. 6.3. They range from 4 pb for lower masses down to about 1 pb for higher mass points. The

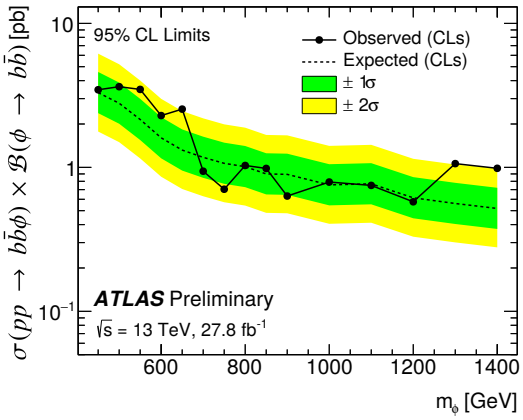


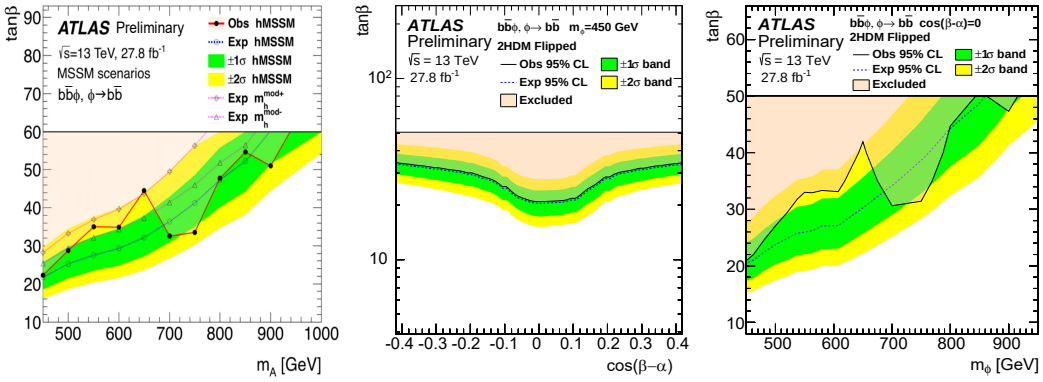
Figure 6.3: Expected and observed upper limits on the $\sigma(pp \to bH/A) \times \mathcal{B}(H/A \to b\bar{b})$ at 95% CL as a function of Higgs boson mass in 27.8 fb^{-1} of pp-collision data at $\sqrt{s} = 13$ TeV.

results of this search can be interpreted in the context of both the MSSM (2HDM type-II) and flipped 2HDM. The MSSM interpretation is performed following specific benchmark scenarios summarised in Ref [202] and the resulting 95% confidence-level limits of $\tan\beta$ versus m_A are shown in Fig. 6.4 left for the hMSSM [203] scenario (with h boson mass of 125 GeV and SUSY partners too heavy for direct detection) together with the expected sensitivities for MSSM variations of so-called m_h^{mod+} and m_h^{mod-} [204]. These limits are comparable to those obtained with the ATLAS $H^{\pm} \to \tau\nu$ search [56] (Sec. 6.3) but weaker than those from $H/A \to \tau\tau$ searches [76, 205] (Sec. 6.2). The middle panel of Fig. 6.4 presents $\tan\beta$ exclusion limits as a function of $\cos(\beta-\alpha)$[2] assuming flipped 2HDM (type-Y) scenario for the $m_{H/A} = 450$ GeV. Right panel of this Figure presents $\tan\beta$ limits as a function of Higgs boson mass $m_{H/A}$ (denoted m_ϕ in the plot) assuming alignment limit $\cos(\beta-\alpha) = 0$. In these interpretations CP conservation in the 2HDM is assumed and the masses of heavy Higgs bosons (both charged and neutral) are equal.

6.1.7 Summary

The presented results show no sign of the existence of an additional heavy Higgs boson. Similar conclusions were derived from the results of similar search in the CMS experiment [207]. That analysis benefited from the full 36 fb^{-1} dataset and used functional description of signal and background shape resulting in lower number of degrees of freedom in the fit and therefore improved sensitivity. On the other hand no categorisation with respect to jet multiplicities was made and no invariant mass rotation was applied there. The achieved sensitivity was nearly identical to that of the described ATLAS analysis. The exclusion limits obtained in that analysis is presented in Fig. 6.5.

[2] α is the neutral CP-even Higgs mixing angle [206]



(a) (b) (c)

Figure 6.4: The 95% CL exclusion limits for MSSM scenarios as a function of m_A (left) and for type-Y 2HDM as functions of $\cos(\beta - \alpha)$ (middle) and for heavy Higgs boson mass denoted m_ϕ (right). Values of $\tan\beta > 60$ are not shown as for such high values the Higgs couplings become non-perturbative [53].

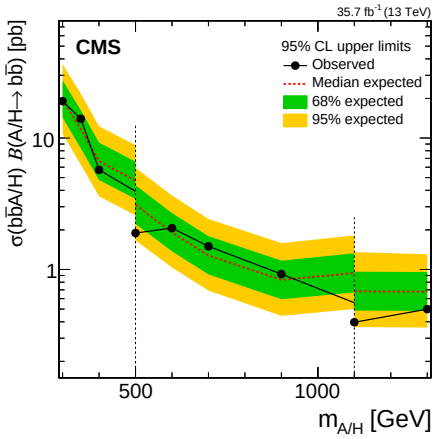


Figure 6.5: CMS results for the expected and observed upper limits on the $\sigma(pp \to bH/A) \times \mathcal{B}(H/A \to b\bar{b})$ at 95% CL as a function of Higgs boson mass [207].

As the presented analysis is limited by data statistics, it will be extended to the full Run-2 dataset of about 140 fb^{-1} and the work is ongoing. In addition, many improvements to the analysis procedure are being currently implemented, including more sophisticated and less conservative approach to background estimation, the use of MVA techniques (including Deep Neural Networks) and, possibly, functional description of signal and background shapes. Publication of results is planned for 2020.

6.2 Search for heavy neutral resonances decaying into pairs of τ leptons

The di-τ decay channel is important for the Higgs sector not only in the Standard Model but also in the 2HDM, especially in its type-II (see Table 2.2) which may be a part of the long-sought Minimal Supersymmetric Standard Model. In this type, heavy Higgs couplings to τ leptons are strongly enhanced at large $\tan\beta$ values. In addition to better signal-to-background ratio compared to the $b\bar{b}$ channel, this makes the di-τ channel crucial for type-II 2HDM searches. Model-independent search for a generic resonance decaying into a pair of τ leptons does not need to be limited in its interpretation to the Higgs sector only. Therefore, in the described search [76] also the Z' bosons produced in Drell-Yan process are considered, especially as many models contain a Z'_{NU} boson that exhibits enhanced couplings to τ leptons [72–74]. A frequently used model is the Sequential Standard Model [65], where Z' boson couplings are identical to those of SM Z boson.

6.2.1 Data and simulation samples

This search uses Run-2 ATLAS data collected in 2015 and 2016 at centre-of-mass energy $\sqrt{s} = 13$ TeV, corresponding to the integrated luminosity of 36 fb^{-1}. Events for analysis are selected using either single-τ triggers with p_{T} thresholds of 80, 125 or 160 GeV, depending on the data-taking period, or single-lepton triggers with lowest p_{T} thresholds ranging from 20 to 26 GeV.

Simulated Higgs signal events were produced with the Powheg-Box and MadGraph5_aMC@NLO+Pythia 8 generators for the ggF and b-associated production modes, respectively. The Z' signal was modelled using Z/γ^* sample enriched in events with high invariant mass from the Pythia 8 generator and correct spin effects in the τ decays were added with TauSpinner-based re-weighting.

The Z/γ^*+jets background was simulated with Powheg-Box interfaced to Pythia 8, similarly to one of the W+jets samples. Another W+jets sample was produced with the Sherpa 2.2.0 generator, whereas top-quark background was modelled using Powheg-Box generator interfaced with Pythia 6 for parton showering and hadronisation. Sherpa 2.1.1 was used to model diboson (ZZ, WZ, WW) backgrounds.

6.2.2 Selection of events and analysis categories

Following the τ trigger selection the events are accepted in the τ_{had}-τ_{had} category if they contain exactly two candidates for hadronic τ decays with $p_{\mathrm{T}} > 65$ GeV each, with no electrons or muons. The leading candidate has to be matched with the online-selected one and exceed its trigger-level p_{T} by 5 GeV. The leading candidate must fulfil *medium* identification criteria whereas the sub-leading one should be identified at *loose* level. In addition, opposite electric charge signs of the two candidates are required and their azimuthal angular separation should be larger than 2.7 as the τ leptons in decays of heavy resonances are produced back-to-back in the transverse plane.

Events selected by the single-lepton triggers are considered for the τ_{lep}-τ_{had} analysis channel. They are required to contain exactly one isolated lepton matched to the triggered object and no additional light leptons. At least one hadronic τ candidate satisfying *medium* identification criteria is required. Only events with opposite signs of electric charges of the lepton and τ candidate, as well as the azimuthal angular separation of the τ and the lepton

above 2.4 are accepted for further studies. Additional suppression of the W+jets background is achieved by requiring the transverse mass of the lepton+$E_{\mathrm{T}}^{\mathrm{miss}}$ system (defined in Sec. 4.1.5) to be no larger than 40 GeV. Additionally, events with the visible mass of the lepton and hadronically-decaying τ within the range of 80 to 110 GeV are rejected to suppress the $Z \to ee$ background.

Events in the two channels are further categorised to exploit the differences in Higgs production modes, into *b-tag* and *b-veto* categories containing at least one b-tagged jet at 70%-efficiency working point and no b-tagged jets, respectively. No categorisation is made for the Z' search.

The invariant mass of the di-τ system would be a good variable separating signal from background. However, due to the presence of neutrinos its reconstruction is very challenging. Therefore, the total transverse mass is used instead. It is defined as $m_{\mathrm{T}}^{\mathrm{tot}} = \sqrt{(p_{\mathrm{T}}^{\tau_1} + p_{\mathrm{T}}^{\tau_2} + E_{\mathrm{T}}^{\mathrm{miss}})^2 - (\mathbf{p}_{\mathrm{T}}^{\tau_1} + \mathbf{p}_{\mathrm{T}}^{\tau_2} + \mathbf{E}_{\mathrm{T}}^{\mathrm{miss}})^2}$ with $\mathbf{p}_{\mathrm{T}}^{\tau_{1,2}}$ being the vectors of momenta of the visible τ decay products (hadrons or leptons) projected onto the transverse plane. Other techniques of invariant mass reconstruction were checked with no significant improvement of the analysis sensitivity.

6.2.3 Estimation of background

The background composition varies between analysis channels and categories. In the τ_{had}-τ_{had} channel the dominant source of background is the multi-jet production. It is estimated using a data-driven technique of fake factors, similar to that described in Sec. 5.2.1 and Ref. [37]. The fake factors are measured in a di-jet-fakes region, in which the events are selected using prescaled single-jet triggers and are required to contain two not-identified back-to-back τ candidates with opposite charges and transverse momenta above 85 and 65 GeV. The sub-leading τ candidate is used for measuring the fake factors. These are then used to multiply the number of multi-jet events obtained from a control region defined as the signal region with the exception of τ identification which has to be failed at *loose* level by the sub-leading candidate, with non-multi-jet backgrounds subtracted based on simulations.

Other important background sources in the τ_{had}-τ_{had} channel are W+jets in the b-veto and $t\bar{t}$ in the b-tag categories. In all the processes the hadronic τ decays are mimicked by jets so, instead of applying τ identification criteria to the events, they are weighted by *fake rates*, defined as the numbers of candidates passing identification to the total number of candidates, and estimated from control regions. These are defined by requiring a single muon with $p_{\mathrm{T}} > 55$ GeV, no electrons and a hadronic τ candidate with p_{T} above 50 GeV, back-to-back with respect to the muon. Transverse mass of the $E_{\mathrm{T}}^{\mathrm{miss}}$ +lepton system should be above 40 GeV. The events in the b-tag category enter top control region whereas those in b-veto category are used for the W control region.

In addition, the background contributions of the $Z/\gamma^* \to \tau\tau$, diboson and $Z/\gamma^* \to \ell\ell$+jets are estimated from simulation.

In the τ_{lep}-τ_{had} channel the dominant background also arises from events in which a jet is mimicking a hadronically-decaying τ lepton. These contributions are estimated similarly to the multi-jet contribution to the τ_{had}-τ_{had} channel. Two control regions for multi-jet estimate are constructed by requiring the hadronic τ candidate to fail the identification criteria (CR1) and by inverting the lepton isolation requirement in addition (CR2). Non-multi-jet contribution in CR2 is subtracted using simulated events and then weighted by

lepton-isolation fake factor, measured in a region with $m_{\mathrm{T}} < 30$ GeV, one lepton and no hadronic τ candidates, yielding the multi-jet contribution in CR1. It is scaled by multi-jet fake factor measured in another control region, obtained by removing any hadronic τ identification requirement and requiring the lepton to fail isolation criteria, and used in the signal region. After subtracting multi-jet part from CR1 the remaining events are used to model W+jets ($t\bar{t}$) backgrounds in the b-veto (b-tag) category (with other processes subtracted based on simulations). Another correction factor is applied to these contributions to estimate their yields in the signal region. This factor is measured in a W-enriched region of high m_{T} in addition to CR1 criteria.

6.2.4 Systematic uncertainties

The systematic uncertainties affecting the precision of the presented search originate from many sources. The simulation samples are affected by the efficiencies of reconstruction, identification and triggering of the physics objects used in the analysis. These include uncertainties on determination of energy scale and identification efficiency of hadronically-decaying τ leptons (both online and offline), electron identification efficiency and small inputs from muon, $E_{\mathrm{T}}^{\mathrm{miss}}$ and jet-related uncertainties. Also, the background processes estimated from simulations are subject to theoretical uncertainties on their cross-section and acceptance determination and are estimated with auxiliary samples produced with different MC generators. Signal samples are also affected by the acceptance uncertainties. Another class of systematic uncertainties affecting this search is related to data-driven estimates of backgrounds. It includes limited statistics of the control regions and a potential mis-modelling of simulation-based subtractions.

6.2.5 Statistical analysis and results

Maximum-likelihood fit to the $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions is performed simultaneously in all analysis channels and categories (b-tag and b-veto categories are merged for Z' search) as well as in the control regions. Signal strength parameter, μ, is the fit parameter of interest and is defined as the ratio of the observed signal to the amount predicted by MSSM or Z' benchmark scenarios. Nuisance parameters are used to parametrise systematic uncertainties. An agreement between the data and predicted background yields is found. The $m_{\mathrm{T}}^{\mathrm{tot}}$ distributions after the fit in the two channels and categories are presented in Fig. 6.6 with various signal hypotheses overlaid.

Following the fit results consistent with the absence of signal, upper exclusion limits on the cross-section times branching fraction for the H/A and Z' bosons are derived at 95% confidence level. The signal strength extracted from the fit is multiplied by predicted cross-section in the MSSM (SSM) for the H/A (Z') case. The limits cover the mass ranges of 0.2-2.25 TeV for the H/A case and up to 4 TeV for the Z' case. Figure 6.7 presents the limits for the ggF and b-associated Higgs production as well as for Z'.

Interpretation of these results is made in MSSM. Left and right panels of Figure 6.8 present the exclusion limits in the $m_{H/A}-\tan\beta$ plane for the $m_{\mathrm{h}}^{\mathrm{mod+}}$ and hMSSM scenarios, respectively. The limits obtained for the latter scenario are stronger because in the $m_{\mathrm{h}}^{\mathrm{mod+}}$ the branching fraction of $H/A \to \tau\tau$ is reduced by the presence of low-mass neutralinos.

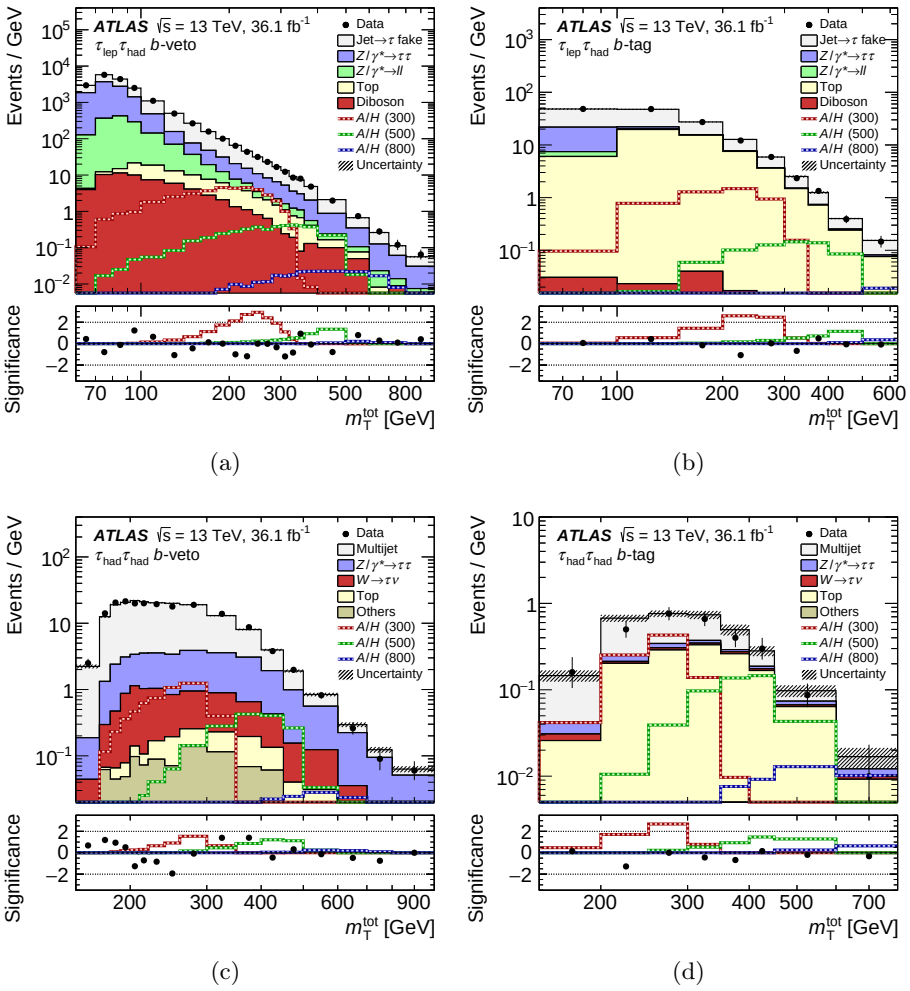
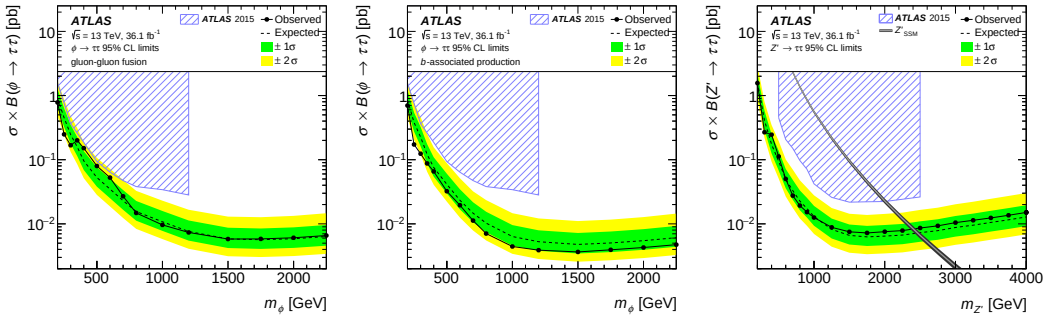


Figure 6.6: Post-fit distribution of $m_{\mathrm{T}}^{\mathrm{tot}}$ in the τ_{lep}-τ_{had} (top row) and τ_{had}-τ_{had} (bottom row) channels, for the b-veto (left column) and b-tag (right column) categories. Signal hypotheses for masses of $m_{H/A} = 300$, 500 and 800 GeV at $\tan\beta = 10$ in the hMSSM scenario are overlaid [76].

6.2.6 Summary

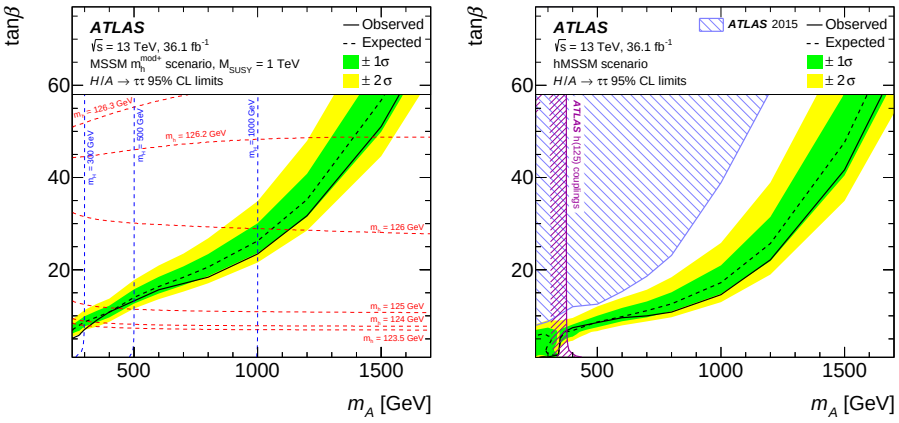
No indication of any excess over the expected SM background was found in this search. Upper limits for cross-section times branching fraction for the searched processes were set at 95% CL. They are in range of 0.78 - 0.0058 pb (0.70 - 0.0037 pb) for gluon-gluon fusion (b-associated) production of H/A bosons for the mass range of 0.2-2.25 TeV and between 1.56 and 0.0072 pb for the Drell-Yan heavy gauge Z' boson for masses between 0.2 and 4 TeV.

Similar search was also performed by the CMS collaboration [205]. The exclusion limits on the heavy neutral Higgs boson production times the $H \to \tau\tau$ branching fraction are obtained for a broader mass range than in the ATLAS search and range from 18 pb (15 pb) at 90 GeV to 0.0035 pb (0.0025 pb) at 3.2 TeV for the ggF (b-associated) production process. In the mass range overlapping with that of ATLAS search the precision is nearly



(a) (b) (c)

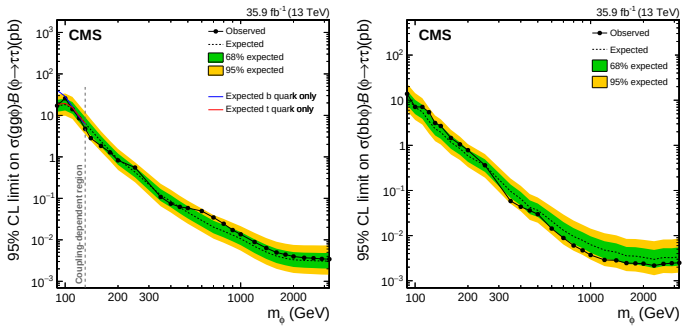
Figure 6.7: The observed and expected 95% CL upper limits on the production cross-section times the branching fraction for H/A (denoted here as ϕ) produced in gluon-gluon fusion (left), b-associated production (middle) and for the Z' boson with predicted SSM Z' cross-section (right) [76]. The hatched area illustrated regions excluded by the previous ATLAS search with 2015 dataset [208].



(a) (b)

Figure 6.8: The observed and expected 95% CL upper limits on $\tan\beta$ as a function of the m_A for the $m_{\mathrm{h}}^{\mathrm{mod+}}$ (left) and hMSSM (right) scenarios. For the hMSSM the previous ATLAS limits [208] are shown as well as the exclusion arising from the SM Higgs couplings measurement [209]. From [76].

identical. The exclusion limits obtained in that analysis for the ggF and b-associated production modes are presented in Fig. 6.9.



(a) (b)

Figure 6.9: The observed and expected 95% CL upper limits on the production cross-section times the branching fraction for H/A (denoted here as ϕ) produced in gluon-gluon fusion (left) and b-associated production (right) in the CMS experiment [205].

6.3 Search for charged Higgs boson in decays into $\tau\nu$ pairs

In addition to the searches for neutral heavy Higgs bosons in di-τ and di-b final states, also the charged Higgs bosons are being looked for in the LHC experiments. This Section presents the ATLAS results of a search for the charged Higgs boson decaying into $\tau\nu$ pair [56]. This decay channel dominates for charged Higgs masses below top quark mass although $H^+ \to cs$ and $H^+ \to cb$ can also be significant for low $\tan\beta$. At higher masses $H^+ \to tb$ decay dominates, but $H^+ \to \tau\nu$ branching fraction can reach 10-15% at higher $\tan\beta$ values. The production channels were described in Sec. 2.1.4 and illustrated in Fig. 2.5. This search probes the type-II 2HDM and the results are interpreted in terms of the MSSM, although the search itself is performed in a model-independent way.

6.3.1 Data and simulation samples

This search bases on the 2015-2016 dataset of $\sqrt{s} = 13$ TeV corresponding to the integrated luminosity of 36 fb^{-1}. Trigger-based selection of sample for analysis uses $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers with thresholds of 70, 90 or 110 GeV, depending on data-taking period, and single-lepton triggers with 24-26 GeV thresholds.

Signal events are simulated with the MadGraph5_aMC@NLO interfaced with Pythia 8 for the three distinct mass regions, below and above top-quark mass as well as in the intermediate region of 160-180 GeV in which the interference between production diagrams is taken into account. Powheg-Box+Pythia 6 generator setup is used to simulate the background arising from top-quark decays, whereas the backgrounds involving W and Z bosons are produced using Sherpa 2.2.1. Diboson backgrounds are simulated using Powheg-Box+Pythia 8 generators.

6.3.2 Selection of events and analysis strategy

The described search targets events with top quarks produced in association with the charged Higgs boson. Analysis channels are defined based on whether the top quark decays semi-leptonically or hadronically, as τ_{had} +lepton or τ_{had} +jets, respectively. In the former, the selection criteria are set as follows. After single-lepton trigger selection exactly one offline lepton matched to the trigger one with $p_{\mathrm{T}} > 30$ GeV is required. Depending on the lepton flavour, two sub-channels are defined: τ_{had} +electron and τ_{had} +muon. Exactly one *medium*-identified τ_{had} candidate with p_{T} above 30 GeV and with electric charge sign opposite to that of the lepton is required in addition. Additional selection criteria include the presence of at least one b-tagged jet with $p_{\mathrm{T}} > 25$ GeV followed by the $E_{\mathrm{T}}^{\mathrm{miss}}$ value above 50 GeV. In the τ_{had} +jets analysis channel, following $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger selection, at least one τ_{had} candidate with p_{T} over 40 GeV and satisfying *medium* identification criteria is required and no *loose* leptons with $p_{\mathrm{T}} > 20$ GeV are allowed. Three additional jets with p_{T} above 25 GeV must be present in the selected event with at least one being b-tagged. The offline value of $E_{\mathrm{T}}^{\mathrm{miss}}$ must exceed 150 GeV and the transverse mass of the τ_{had} +$E_{\mathrm{T}}^{\mathrm{miss}}$ system should be above 50 GeV.

The selected events are afterwards used in the actual search for the charged Higgs boson signal. It is performed with a discriminating variable constructed using multivariate analysis techniques, namely the BDT algorithms. The discriminating variables are constructed separately for the τ_{had} +lepton and τ_{had} +jets channel and are based on kinematic properties of the analysed objects. Input variables include $E_{\mathrm{T}}^{\mathrm{miss}}$, τ transverse momenta, p_{T}

of b-jets and variables describing angular separation of particular analysis objects (b jets, τ leptons, light leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$). Last but not least, the Υ variable, as defined in [154] and Sec. 5.1.3 is used for events with single-prong τ_{had} candidates, significantly increasing the separation between signal and background events for lower H^+ masses. The training of the BDT algorithms is performed separately in several H^+ mass regions to increase the search sensitivity.

6.3.3 Background estimation

The backgrounds for this analysis involving the decays of W and Z bosons are estimated directly from simulations for cases where the reconstructed hadronically-decaying τ candidate originates from a genuine τ decay or from light lepton. The $t\bar{t}$ background normalisation is derived from a fit to data. The multi-jet background is estimated with a data-driven method. It involves the computation of *fake factors*, following similar principles to the SM $H \to \tau\tau$ analyses (see Sec. 5.2) and $H/A/Z' \to \tau\tau$ search (see Sec. 6.2). In this search the fake factors for a given control region are defined as the ratio of the number of jets reconstructed as τ_{had} candidates that pass the identification selection to the number of jets mimicking τ_{had} candidates and failing the *medium* τ identification criteria while satisfying the *loose*. The fake factors are measured in regions with different fractions of quark- and gluon-initiated jets defined by signal-selection criteria with a b-jet veto and $E_{\mathrm{T}}^{\mathrm{miss}} < 80$ GeV requirement for the former and with b-jet veto, no $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement and the lepton-$E_{\mathrm{T}}^{\mathrm{miss}}$ transverse mass between 60 and 160 GeV for the latter. The fractions of the two background types are then estimated with a template fit in the control regions to variables sensitive to jet-type composition and combined fake factors are derived.

The data-driven multi-jet background estimate cannot however predict the shape of the Υ variable in the signal region as the τ identification input variables are strongly correlated with Υ. On the other hand, the Υ is not correlated with any other of the analysis BDT input variables and an inverse transform sampling method [210] can be used. It employs cumulative distribution functions F measured in the background control (CR) and signal regions (SR) to obtain the shape of background Υ distribution in the SR: $\Upsilon_{\mathrm{SR}}^{\mathrm{multijet}} = F_{\mathrm{SR}}^{-1}(F_{\mathrm{CR}}(\Upsilon))$.

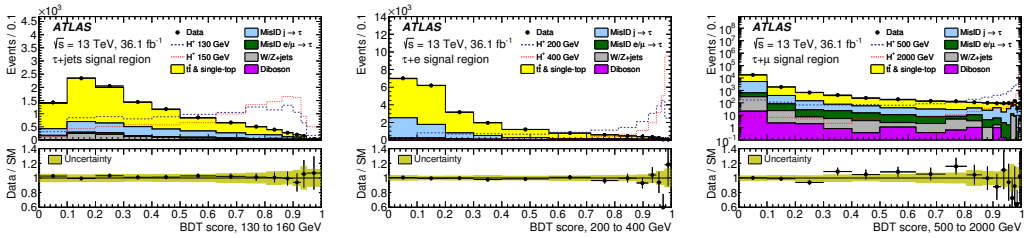
The data-driven background estimation procedure is eventually validated by comparing the predicted and measured BDT distributions in regions enriched in either $t\bar{t}$ or multi-jet backgrounds and depleted with signal, constructed by requiring an $e\mu$ pair (instead of e or μ accompanying τ_{had}) or by vetoing the presence of b-jets. The predicted BDT score distributions were found to agree with measured ones.

6.3.4 Systematic uncertainties

Systematic uncertainties affecting the shapes and normalisations of signal and background BDT score distributions arise from several sources. The most important one is the modelling of backgrounds with the fake-factor method. It is related to the uncertainty on quark- and gluon-initiated jet fractions in the control regions, contamination from genuine τ_{had} and limited statistics of the control regions. In addition, identification efficiencies for τ leptons, determination of the τ and jet energy scale (and its impact on $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation) and b-tagging efficiencies are also considered. Moreover, theoretical uncertainties on the modelling of signal, $t\bar{t}$ and electroweak-boson backgrounds are taken into account.

6.3.5 Results

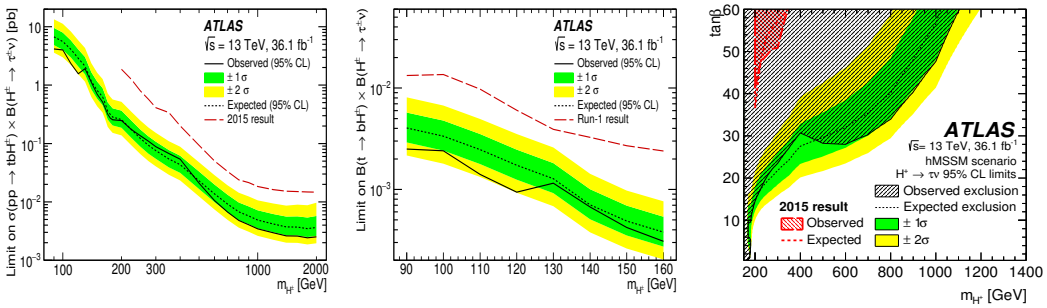
The signal cross-sections times branching fraction, $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ in the whole considered mass range and $\mathcal{B}(t \to bH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ for lower masses are extracted from data using a maximum-likelihood fit, performed simultaneously in three signal regions ($\tau_{\text{had}} + e/\mu$, τ_{had} +jets) and in $t\bar{t}$ control region. The results are consistent with the absence of signal. Example distributions of BDT scores in the three signal regions and for various mass ranges after the fit are presented in Fig. 6.10. Following the fit results,



(a) (b) (c)

Figure 6.10: Post-fit BDT score distributions for τ_{had} +jets channel in the mass range of $130 < m_{H^+} < 160$ GeV (left), τ_{had} +electron for $160 < m_{H^+} < 180$ GeV (middle) and τ_{had} +muon for $500 < m_{H^+} < 2000$ GeV [56].

exclusion limits at 95% CL are set on the $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ in the mass range of 90 - 2000 GeV and on $\mathcal{B}(t \to bH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ for masses below 160 GeV. Interpretation of results is made in the hMSSM context and 95%-CL exclusion limits on $\tan\beta$ as a function of charged Higgs mass are derived. The obtained exclusion limits are shown in Fig. 6.11.



(a) (b) (c)

Figure 6.11: Observed and expected 95% CL exclusion limits on the $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ (left), $\mathcal{B}(t \to bH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ (middle) and on $\tan\beta$ as a function of m_{H^+} in the context of hMSSM (right). The results from early-Run-2 (2015) and Run-1 analyses are shown for comparison [56].

6.3.6 Summary

The results of the described search are consistent with no-signal hypothesis. Exclusion limits are derived for the H^+ production cross-section times the $H^+ \to \tau\nu$ branching fraction in the 90 - 2000 GeV mass range and are between 4.2 and 0.0025 pb. The limits on the $\mathcal{B}(t \to bH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ branching fractions is between 0.25% and 0.031% for the mass range of 90 - 160 GeV. Preliminary results from the CMS experiment [211] were obtained with a dataset of similar integrated luminosity and are presented in the charged Higgs boson mass range between 80 and 3000 GeV and the resulting exclusion limits on $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to \tau\nu)$ are between 6 and 0.005 pb, in agreement with the described ATLAS results. The exclusion limits obtained in that analysis are presented in Fig. 6.12.

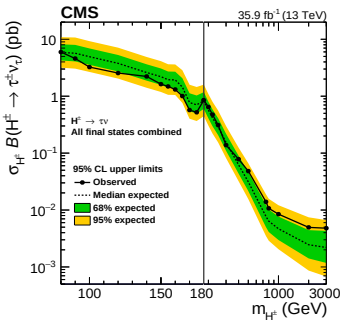


Figure 6.12: Observed and expected 95% CL exclusion limits on the $\sigma_{H^+} \times \mathcal{B}(H^+ \to \tau\nu)$ in the search by the CMS experiment [211].

6.4 Search for charged Higgs boson in decays into tb pairs

For the charged Higgs boson with masses above the top-quark mass threshold the dominant decay mode is $H^+ \to tb$ in broad range of models [34, 202]. Search in this channel performed by the ATLAS experiment [212] is described here. It is complementary to that in the $H^+ \to \tau\nu$ channel presented in Sec. 6.3 and [56]. The dominant production process for the considered charged Higgs boson masses is shown in middle panel of Fig. 2.5.

6.4.1 Samples for analysis

This search uses the 2015-2016 ATLAS dataset of pp collisions collected at $\sqrt{s} = 13$ TeV and corresponding to the integrated luminosity of 36 fb^{-1}. Analysis events were selected with single-lepton triggers with p_{T} thresholds of 20-26 GeV for muons and 24-26 GeV for electrons, depending on the data-taking period.

Monte Carlo samples used to model the expected signal were produced with the MadGraph5_aMC@NLO generator interfaced with Pythia 8 for hadronisation and parton showering. Signal samples for several H^+ mass hypotheses were generated throughout the $200 < m_{H^+} < 2000$ GeV mass range. The $t\bar{t}$ background events were simulated with the Powheg-Box program with Pythia 8 for showering and hadronisation and EvtGen for b- and c-hadron decays. The Sherpa 2.2.1 generator was used for modelling of the W/Z+jets samples and, in addition, the $t\bar{t}W/Z$ events were obtained from the MadGraph5_aMC@NLO+Pythia 8 setup.

6.4.2 Selection of events

The selection procedure in this search targets $H^+ \to tb$ events with either one charged lepton and jets (ℓ+jets channel), or with two charged leptons and jets ($\ell\ell$ channel) in the final state. Following the trigger selection, at least one offline electron or muon, matched to the online-reconstructed candidate and with p_{T} above 27 GeV is required. Additional leptons are required to have $p_{\mathrm{T}} > 10$ GeV or 15 GeV in case of di-electron events. The ℓ+jets and $\ell\ell$ channel selections are defined to be mutually exclusive. In the ℓ+jets channel at least five jets are required, of which at least two have to be b-tagged at 70%-efficiency working point. For the $\ell\ell$ channel, only events with at least three jets, including two b-tagged ones (at 70% efficiency working point), are accepted. In addition, in the ee and $\mu\mu$ channels the di-lepton invariant mas should be above 15 GeV and not within the Z-boson mass window of 83-99 GeV.

In order to increase the analysis sensitivity, the events passing signal selection are further categorised into signal regions (SR), with enhanced possible signal contribution, and control regions (CR). The categorisation is based on the number of jets (j) and number of b-tagged jets (b). Four SRs (5j3b, 5j≥4b, ≥6j3b and ≥6j≥4b) and two CRs (5j2b and ≥6j2b) are defined for the ℓ+jets channel, whereas in the $\ell\ell$ channel there are two SRs (≥4j3b and ≥4j≥4b) and two CRs (3j2b and ≥4j2b).

6.4.3 Estimation of background

The estimation of the relevant background contributions to the signal regions is estimated depending on the particular background source. Processes with prompt leptons are estimated using simulated samples. The number of $t\bar{t}$ production events is overestimated in

the simulation and needs therefore to be re-weighted. The re-weighting function is obtained from the control region with exactly four jets with at least two of them b-tagged. Another data-driven correction is applied to the simulated Z+HF contribution. It is extracted from a region with two oppositely-charged same-flavour leptons with invariant mass in the range of $83 < m_{\ell\ell} < 99$ GeV, compatible with that of Z boson.

The selection criteria can also be satisfied by events containing non-prompt leptons, from semi-leptonic hadron decays or jets from multi-jet production misidentified as leptons. In the ℓ+jets channel a matrix method [213] is used. Event sample containing leptons is selected using criteria with relaxed isolation or lepton identification requirements and then weighted based on the efficiencies for prompt and non-prompt leptons. The efficiencies are determined from data in dedicated control regions. In the $\ell\ell$ channel this contribution is modelled with simulations and the normalisation is extracted by comparing data to simulations in a region with same-sign di-lepton events. Multi-jet contribution is negligible in this channel.

6.4.4 Multivariate analysis

In each of the signal regions a BDT discriminant is defined and trained against all backgrounds (ℓ+jets channel) or against $t\bar{t}$ background ($\ell\ell$ channel). The most important input variables in the ℓ+jets channel for H^+ masses above 400 GeV are the scalar sum of jet transverse momenta, $H_{\mathrm{T}}^{\mathrm{jets}}$ and the p_{T} of the leading jet. At masses below 300 GeV a special kinematic discriminant described below as well as the invariant mass of the non-b-tagged jet pair with smallest angular separation have the largest impact. The aforementioned kinematic discriminant reflects the probability that the event is compatible with the $H^+ \to tb$ or $t\bar{t}$ decays and is constructed as product of probability density functions for the invariant masses of semi-leptonically-decaying top quark, hadronically-decaying W boson, their mass difference and the difference between the charged Higgs mass (or tb mass with b-jet not associated to top decay in case of $t\bar{t}$ background) and the mass of hadronically or leptonically-decaying top quark (depending on the decay channel of the top quark from Higgs decay). The probability density functions are obtained from simulations with the reconstructed objects matched to generator-level ones.

In the $\ell\ell$ channel for charged Higgs masses below 600 GeV the smallest invariant mass formed by two b-tagged jets as well as the smallest invariant mass formed by a lepton and b-tagged jet have the strongest impact on the discriminant power. For the masses above 600 GeV the most important variables are the scalar sum of all jet- and lepton-p_{T}, $H_{\mathrm{T}}^{\mathrm{all}}$ and the transverse momentum of the jet pair with maximum p_{T}.

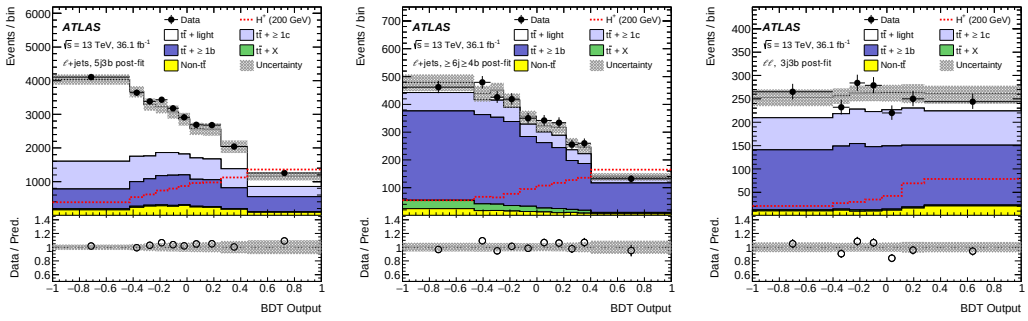
6.4.5 Systematic uncertainties

Various sources of systematic uncertainty can affect the shapes and normalisations of the final BDT distributions. The most important sources include flavour tagging efficiency scale factors, jet energy scale and resolution (and its impact on $E_{\mathrm{T}}^{\mathrm{miss}}$), modelling of the $t\bar{t}$ background and the limited statistics of simulation samples available for this search. The modelling of the expected signal process has a negligible impact.

6.4.6 Statistical analysis and results

A binned maximum-likelihood fit to data is performed simultaneously in all analysis categories. Each mass hypothesis is tested separately. The inputs to the fit include the event yields in CRs and the binned BDT score distributions in the SRs. The parameter of interest is the cross-section times branching fraction, $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to tb)$. Nuisance parameters parametrising the systematic uncertainties and the $t\bar{t}$ background normalisation factors, as well as statistical uncertainties are included in the fit.

The results obtained with the fit are consistent with the absence of signal. Example distributions of the BDT scores after the fit for the charged Higgs mass hypothesis of 200 GeV are shown in Fig. 6.13.



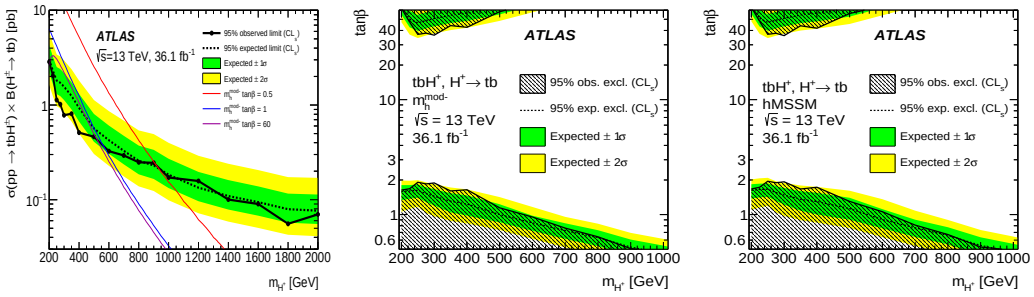
(a) (b) (c)

Figure 6.13: Post-fit BDT score distributions for ℓ+jets channel and 5j3b region (left), ℓ+jets ≥6j≥4b (middle) and $\ell\ell$ 3j3b (right) for charged Higgs mas hypothesis m_{H^+} = 200 GeV [212].

The 95% CL exclusion limits for the $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to tb)$ are derived. They range from 2.9 pb at $m_{H^+} = 200$ GeV to 0.07 pb at $m_{H^+} = 2000$ GeV. The results are also interpreted in the $m_h^{\text{mod}-}$ scenario of the MSSM and the hMSSM. The limit plots are presented in Fig. 6.14.

6.4.7 Summary

The presented search resulted in no hints for the existence of the charged Higgs boson in the mass range of 200 - 2000 GeV. The exclusion limits in the $\tan\beta - m_{H^+}$ plane are complementary to those of the $H^+ \to \tau\nu$ search described in Sec. 6.3 and Ref. [56] and generally cover more of the low-$\tan\beta$ region for masses below 1 TeV, whereas in the high-$\tan\beta$ region the search in $\tau\nu$ final state the limits are stronger. No Run-2 result in the tb channel from the CMS collaboration was available by the time of writing of this monograph.



(a) (b) (c)

Figure 6.14: Observed and expected 95% CL exclusion limits on the $\sigma(pp \to tbH^+) \times \mathcal{B}(H^+ \to tb)$ with $m_h^{\mathrm{mod}-}$ theory predictions overlaid (left) and on $\tan\beta$ as a function of the charged Higgs mass m_{H^+} in the $m_h^{\mathrm{mod}-}$ (middle) and hMSSM (right) scenarios [212].

6.5 Summary

The presented results of Beyond-Standard-Model Higgs-boson searches show no significant deviation from the Standard-Model predictions. Exclusion limits, narrowing the possible search regions for several variants of New Physics models are derived. There are still several unexplored and not excluded regions in the parameter spaces of the considered BSM models. The analyses covering the full Run-2 dataset are ongoing and their results will soon supersede those presented in this Chapter.

Chapter 7

Summary

The successful operation of the LHC in the past 9 years yielded a great set of physics results. This was preceded by an enormous amount of work during long years of planning, constructing and testing both the accelerator and the large and complex particle detectors. Even the 2008 LHC accident did not eventually prevent the physicists from exploring the previously inaccessible energy regimes.

The results presented in this monograph could cover only a part of ATLAS results with third-generation fermions in final states. This is a tiny yet important subset of all results presented by the ATLAS collaboration during its up-to-now operation. Measurements in the Standard Model area, after 2012 including the 125-GeV Higgs sector, are still crucial for understanding the nature of our Universe. Without them, the searches for Beyond Standard Model phenomena would not yield as significant and stringent exclusion limits, that narrow the field for future searches. All in all, no precise physics results would be possible to obtain without a good understanding of the detector and of the reconstruction and identification of individual physics objects in each event. This is achieved by constant work on development of the reconstruction and identification procedures and was also discussed in this monograph.

At the current moment the results of analyses based on the early Run-2 data are available. They already supersede Run-1 results in the search category, not only because of higher collision energies (and, correspondingly, larger expected New Physics cross-sections) but also due to larger amount of data available. These results will soon be superseded again by those obtained with the full Run-2 dataset, corresponding to the integrated luminosity of about 149 fb^{-1} (four times more than the early-Run-2 dataset), and with new or refined experimental techniques, such as mass regression, improved precision in the determination of data-to-simulation scale factors and more.

Following the discovery of the Higgs boson in 2012 the Standard Model seems complete. This is also supported by the results presented in this monograph concerning the Standard-Model scalar sector that is well understood and, so far, no deviations from theoretical predictions are observed. Moreover, no signs of New Physics appear in the direct searches for new particles, including the described analyses in the extended Higgs sector.

With not a single hint for New Physics at the LHC it may well be possible that it will not be reachable by this accelerator, even after its upgrade to the High-Luminosity LHC. Even the planned construction of the Future Circular Collider may not necessarily provide access to New Physics if it simply lies too far above the reachable energy thresholds. It does not mean that the particle colliders would not be needed in the near future -

better and better understanding of the known processes is always of high value. It might however be, that the discoveries will rather happen in non-collider experiments, such as those oriented for astroparticle observations, such as searches for Dark Matter, high-energy neutrino investigations or observations of air showers.

Bibliography

[1] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964) 214, ISSN: 0031-9163, URL: http://www.sciencedirect.com/science/article/pii/S0031916364920013.

[2] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 1*, tech. rep. CERN-TH-401, CERN, 1964, URL: http://cds.cern.ch/record/352337.

[3] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 2*, (1964) 80 p, Version 1 is CERN preprint 8182/TH.401, Jan. 17, 1964, URL: http://cds.cern.ch/record/570209.

[4] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.

[5] M. L. Perl et al., *Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation*, Phys. Rev. Lett. **35** (22 1975) 1489, URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[6] G. J. Feldman, *e+e- Annihilation*, Proceedings of the 19th International Conference on High Energy Physics, Tokyo, August 23-30, 1978 (1978) 777.

[7] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (5 1977) 252, URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[8] Belle Collaboration, *Measurement of the Moments of the Photon Energy Spectrum in B —> X(s) gamma Decays and Determination of |V(cb)| and m(b) at Belle*, Phys. Rev. **D78** (2008) 032016, arXiv: 0803.2158 [hep-ex].

[9] H. A. Y. Jin D. Epifanov, *Latest Belle results on Tau decays*, Proceedings for the 15th International Workshop on Tau Lepton Physics, Amsterdam, The Netherlands, 24-28 September 2018 (2018).

[10] BaBar Collaboration, *Measurement and interpretation of moments in inclusive semileptonic decays anti-B —> X(c) l- anti-nu*, Phys. Rev. **D81** (2010) 032003, arXiv: 0908.0415 [hep-ex].

[11] A. Lusiani, *Measurement of the Tau Lepton Lifetime with BABAR*, Nuclear Physics B Proceedings Supplements **144** (2005) 105.

[12] The DELPHI Collaboration, *Study of b-quark mass effects in multijet topologies with the DELPHI detector at LEP*, Eur. Phys. J. C **55** (2008) 525, URL: https://doi.org/10.1140/epjc/s10052-008-0631-5.

[13] The DELPHI Collaboration, *A precise measurement of the tau lifetime*, Eur. Phys. J. C **36** (2004) 283, URL: https://doi.org/10.1140/epjc/s2004-01953-7.

[14] TeVI Group, *Design Report Tevatron 1 project*, tech. rep. FERMILAB-DESIGN-1984-01, 1984, URL: https://cds.cern.ch/record/1478620.

[15] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001, URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[16] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001, URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[17] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: 1106.0034 [hep-ph].

[18] J. F. Gunion, S. Dawson, H. E. Haber, and G. L. Kane, *The Higgs hunter's guide*, vol. 80, Brookhaven Nat. Lab., 1989, URL: https://cds.cern.ch/record/425736.

[19] D0 Collaboration, *Observation of the Top Quark*, Phys. Rev. Lett. **74** (14 1995) 2632, URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2632.

[20] CDF Collaboration, *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (14 1995) 2626, URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[21] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results*, Submitted to: Eur. Phys. J. (2018), arXiv: 1810.01772 [hep-ex].

[22] CMS Collaboration, *Measurement of the top quark mass with lepton+jets final states using* p p *collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. **C78** (2018) 891, arXiv: 1805.01428 [hep-ex].

[23] DONUT Collaboration, *Observation of tau neutrino interactions*, Phys. Lett. **B504** (2001) 218, arXiv: hep-ex/0012035 [hep-ex].

[24] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[25] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. **D69** (2004) 094008, arXiv: hep-ph/0312266 [hep-ph].

[26] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, arXiv: 0901.0002 [hep-ph].

[27] ATLAS Collaboration,
Measurement of the $W \to \ell\nu$ and $Z/\gamma^* \to \ell\ell$ production cross sections in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector,
JHEP **12** (2010) 060, arXiv: 1010.2130 [hep-ex].

[28] ATLAS Collaboration, *Measurement of the $W \to \tau\nu_\tau$ cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment*, Phys. Lett. B **706** (2012) 276,
arXiv: 1108.4101 [hep-ex].

[29] CMS Collaboration, *The CMS experiment at the CERN LHC*,
JINST **3** (2008) S08004.

[30] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*,
Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex].

[31] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30,
arXiv: 1207.7235 [hep-ex].

[32] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.),
Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,
CERN-2011-002 (CERN, Geneva, 2011), arXiv: 1101.0593 [hep-ph].

[33] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.),
Handbook of LHC Higgs Cross Sections: 2. Differential Distributions,
CERN-2012-002 (CERN, Geneva, 2012), arXiv: 1201.3084 [hep-ph].

[34] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.),
Handbook of LHC Higgs Cross Sections: 3. Higgs Properties,
CERN-2013-004 (CERN, Geneva, 2013), arXiv: 1307.1347 [hep-ph].

[35] S. Alioli, P. Nason, C. Oleari, and E. Re,
NLO Higgs boson production via gluon fusion matched with shower in POWHEG,
JHEP **04** (2009) 002, arXiv: 0812.0578 [hep-ph].

[36] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02** (2010) 037,
arXiv: 0911.5299 [hep-ph].

[37] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **04** (2015) 117,
arXiv: 1501.04943 [hep-ex].

[38] CMS Collaboration,
Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons,
JHEP **05** (2014) 104, arXiv: 1401.5041 [hep-ex].

[39] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*,
JHEP **08** (2016) 045, arXiv: 1606.02266 [hep-ex].

[40] F. Englert and R. Brout,
Broken Symmetry and the Mass of Gauge Vector Mesons,
Phys. Rev. Lett. **13** (1964) 321.

[41] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*,
Phys. Lett. **12** (1964) 132.

[42] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*,
Phys. Rev. Lett. **13** (1964) 508.

[43] H. Miyazawa, *Baryon Number Changing Currents**,
Progress of Theoretical Physics **36** (1966) 1266, eprint: /oup/backfile/content_public/journal/ptp/36/6/10.1143/ptp.36.1266/2/36-6-1266.pdf,
URL: http://dx.doi.org/10.1143/PTP.36.1266.

[44] H. Miyazawa, *Spinor Currents and Symmetries of Baryons and Mesons*,
Phys. Rev. **170** (5 1968) 1586,
URL: https://link.aps.org/doi/10.1103/PhysRev.170.1586.

[45] J.-L. Gervais and B. Sakita,
Field theory interpretation of supergauges in dual models,
Nuclear Physics B **34** (1971) 632, ISSN: 0550-3213, URL:
http://www.sciencedirect.com/science/article/pii/0550321371903518.

[46] P. Ramond, *Dual Theory for Free Fermions*, Phys. Rev. D **3** (10 1971) 2415,
URL: https://link.aps.org/doi/10.1103/PhysRevD.3.2415.

[47] J. Wess and B. Zumino, *Supergauge transformations in four dimensions*,
Nuclear Physics B **70** (1974) 39, ISSN: 0550-3213, URL:
http://www.sciencedirect.com/science/article/pii/0550321374903551.

[48] P. Fayet, *Supersymmetry and weak, electromagnetic and strong interactions*,
Physics Letters B **64** (1976) 159.

[49] P. Fayet, *Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions*, Phys. Lett. **69B** (1977) 489.

[50] G. R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*,
Phys. Lett. **76B** (1978) 575.

[51] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*,
Nucl. Phys. **B193** (1981) 150.

[52] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model*, Phys. Rept. **459** (2008) 1,
arXiv: hep-ph/0503173 [hep-ph].

[53] ATLAS Collaboration,
Search for Heavy Neutral Higgs Bosons Produced in Association with b-quarks and Decaying to b-quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector,
ATLAS-CONF-2019-010, 2019, URL: http://cdsweb.cern.ch/record/2669403.

[54] DELPHI, OPAL, ALEPH, LEP Working Group for Higgs Boson Searches, L3,
Search for neutral MSSM Higgs bosons at LEP, Eur. Phys. J. **C47** (2006) 547,
arXiv: hep-ex/0602042 [hep-ex].

[55] CDF and D0 Collaborations, *Combined CDF and D0 Upper Limits on MSSM Higgs Boson Production in tau-tau Final States with up to 2.2 fb-1*, (2010), arXiv: 1003.3363 [hep-ex].

[56] ATLAS Collaboration, *Search for charged Higgs bosons decaying via $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ in the τ+jets and τ+lepton final states with 36 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, JHEP **09** (2018) 139, arXiv: 1807.07915 [hep-ex].

[57] OPAL, DELPHI, L3, ALEPH, LEP Higgs Working Group for Higgs boson searches, "Search for charged Higgs bosons: Preliminary combined results using LEP data collected at energies up to 209-GeV", *Lepton and photon interactions at high energies. Proceedings, 20th International Symposium, LP 2001, Rome, Italy, July 23-28, 2001*, 2001, arXiv: hep-ex/0107031 [hep-ex], URL: http://weblib.cern.ch/abstract?CERN-L3-NOTE-2689.

[58] CDF Collaboration, *Search for charged Higgs bosons in decays of top quarks in p anti-p collisions at s**(1/2) = 1.96 TeV*, Phys. Rev. Lett. **103** (2009) 101803, arXiv: 0907.1269 [hep-ex].

[59] D0 Collaboration, *Search for Charged Higgs Bosons in Top Quark Decays*, Phys. Lett. **B682** (2009) 278, arXiv: 0908.1811 [hep-ex].

[60] C. F. Kolda, *Gauge mediated supersymmetry breaking: Introduction, review and update*, Nucl. Phys. Proc. Suppl. **62** (1998) 266, [,266(1997)], arXiv: hep-ph/9707450 [hep-ph].

[61] OPAL Collaboration, *Searches for gauge-mediated supersymmetry breaking topologies in e+ e- collisions at LEP2*, Eur. Phys. J. **C46** (2006) 307, arXiv: hep-ex/0507048 [hep-ex].

[62] D0 Collaboration, *Search for squark production in events with jets, hadronically decaying tau leptons and missing transverse energy at s=1.96 TeV*, Physics Letters B **680** (2009) 24, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/S0370269309009162.

[63] ATLAS Collaboration, *Search for squarks and gluinos in final states with hadronically decaying τ-leptons, jets, and missing transverse momentum using pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Submitted to: Phys. Rev. (2018), arXiv: 1808.06358 [hep-ex].

[64] CMS Collaboration, *Search for supersymmetry in events with a τ lepton pair and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **11** (2018) 151, arXiv: 1807.02048 [hep-ex].

[65] G. Altarelli, B. Mele, and M. Ruiz-Altaba, *Searching for new heavy vector bosons in $p\bar{p}$ colliders*, Zeitschrift für Physik C Particles and Fields **45** (1989) 109, ISSN: 1431-5858, URL: https://doi.org/10.1007/BF01556677.

[66] J. L. Hewett and T. G. Rizzo,
Low-energy phenomenology of superstring-inspired E6 models,
Physics Reports **183** (1989) 193, ISSN: 0370-1573, URL:
http://www.sciencedirect.com/science/article/pii/0370157389900719.

[67] M. Cvetic and S. Godfrey, *Discovery and identification of extra gauge bosons*, (1995) 383, arXiv: hep-ph/9504216 [hep-ph].

[68] R. Diener, S. Godfrey, and T. A. W. Martin, *Unravelling an Extra Neutral Gauge Boson at the LHC using Third Generation Fermions*,
Phys. Rev. **D83** (2011) 115008, arXiv: 1006.2845 [hep-ph].

[69] J. Abdallah et al., *Simplified Models for Dark Matter Searches at the LHC*,
Phys. Dark Univ. **9-10** (2015) 8, arXiv: 1506.03116 [hep-ph].

[70] G. Busoni et al., *Recommendations on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter*, (2016), ed. by A. Boveia et al., arXiv: 1603.04156 [hep-ex].

[71] M. Fairbairn, J. Heal, F. Kahlhoefer, and P. Tunney, *Constraints on Z' models from LHC dijet searches and implications for dark matter*, JHEP **09** (2016) 018, arXiv: 1605.07940 [hep-ph].

[72] D. J. Muller and S. Nandi, *Top flavor: A Separate SU(2) for the third family*,
Phys. Lett. **B383** (1996) 345, arXiv: hep-ph/9602390 [hep-ph].

[73] K. Hsieh, K. Schmitz, J.-H. Yu, and C. .-.-P. Yuan,
Global Analysis of General SU(2) x SU(2) x U(1) Models with Precision Data,
Phys. Rev. **D82** (2010) 035011, arXiv: 1003.3482 [hep-ph].

[74] E. Malkawi, T. M. P. Tait, and C. P. Yuan,
A Model of strong flavor dynamics for the top quark, Phys. Lett. **B385** (1996) 304, arXiv: hep-ph/9603349 [hep-ph].

[75] CMS Collaboration, *Search for a W' boson decaying to a τ lepton and a neutrino in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. (2018), arXiv: 1807.11421 [hep-ex].

[76] ATLAS Collaboration,
Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,
JHEP **01** (2018) 055, arXiv: 1709.07242 [hep-ex].

[77] ATLAS Collaboration,
A search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,
Phys. Rev. Lett. (2018), arXiv: 1808.00336 [hep-ex].

[78] CMS Collaboration, *Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV*,
Phys. Lett. B **778** (2018) 101, arXiv: 1707.02909 [hep-ex].

[79] ATLAS Collaboration,
Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector,
Phys. Lett. B **786** (2018) 59, arXiv: 1808.08238 [hep-ex].

[80] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, Phys. Rev. Lett. **121** (2018) 121801, arXiv: 1808.08242 [hep-ex].

[81] R. Harlander, M. Kramer, and M. Schumacher, *Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach*, (2011), arXiv: 1112.3478 [hep-ph].

[82] CDF and D0 Collaborations, *Search for Neutral Higgs Bosons in Events with Multiple Bottom Quarks at the Tevatron*, Phys. Rev. **D86** (2012) 091101, arXiv: 1207.2757 [hep-ex].

[83] J. Alwall, P. Schuster, and N. Toro, *Simplified Models for a First Characterization of New Physics at the LHC*, Phys. Rev. **D79** (2009) 075020, arXiv: 0810.3921 [hep-ph].

[84] D. Alves, *Simplified Models for LHC New Physics Searches*, J. Phys. **G39** (2012) 105005, ed. by N. Arkani-Hamed et al., arXiv: 1105.2838 [hep-ph].

[85] ATLAS Collaboration, *Search for $B-L$ R-parity-violating top squarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS experiment*, Phys. Rev. D **97** (2018) 032003, arXiv: 1710.05544 [hep-ex].

[86] ATLAS Collaboration, *Search for supersymmetry in events with b-tagged jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **11** (2017) 195, arXiv: 1708.09266 [hep-ex].

[87] ATLAS Collaboration, *Search for supersymmetry in final states with missing transverse momentum and multiple b-jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **06** (2018) 107, arXiv: 1711.01901 [hep-ex].

[88] CMS Collaboration, *Search for R-parity violating supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV using b jets in a final state with a single lepton, many jets, and high sum of large-radius jet masses*, Phys. Lett. B **783** (2018) 114, arXiv: 1712.08920 [hep-ex].

[89] CMS Collaboration, *Search for top squarks decaying via four-body or chargino-mediated modes in single-lepton final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **09** (2018) 065, arXiv: 1805.05784 [hep-ex].

[90] T. Jacques et al., *Complementarity of DM searches in a consistent simplified model: the case of Z'*, JHEP **10** (2016) 071, arXiv: 1605.06513 [hep-ph].

[91] Y. G. Kim, K. Y. Lee, C. B. Park, and S. Shin, *Secluded singlet fermionic dark matter driven by the Fermi gamma-ray excess*, Phys. Rev. **D93** (2016) 075023, arXiv: 1601.05089 [hep-ph].

[92] ATLAS Collaboration, *Search for resonances in the mass distribution of jet pairs with one or two jets identified as b-jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **98** (2018) 032016, arXiv: 1805.09299 [hep-ex].

[93] CMS Collaboration, *Search for Narrow Resonances in the b-Tagged Dijet Mass Spectrum in Proton–Proton Collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Rev. Lett. **120** (2018) 201801, arXiv: 1802.06149 [hep-ex].

[94] CMS Collaboration, *Search for low-mass resonances decaying into bottom quark-antiquark pairs in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Submitted to: Phys. Rev. (2018), arXiv: 1810.11822 [hep-ex].

[95] ATLAS Collaboration, *Improved luminosity determination in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **73** (2013) 2518, arXiv: 1302.4393 [hep-ex].

[96] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex].

[97] *ATLAS Luminosity Public Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2, Accessed: 2019-01-07.

[98] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[99] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[100] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, URL: https://cds.cern.ch/record/1291633.

[101] L. Adamczyk et al., *Technical Design Report for the ATLAS Forward Proton Detector*, tech. rep. CERN-LHCC-2015-009. ATLAS-TDR-024, 2015, URL: https://cds.cern.ch/record/2017378.

[102] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[103] *ATLAS Computing: technical design report*, Technical Design Report ATLAS, CERN, 2005, URL: https://cds.cern.ch/record/837738.

[104] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[105] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

[106] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph].

[107] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph].

[108] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph].

[109] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[110] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, Eur. Phys. J. C **77** (2017) 195, arXiv: 1612.01456 [hep-ex].

[111] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data*, ATLAS-CONF-2014-032, 2014, URL: https://cds.cern.ch/record/1706245.

[112] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016, URL: https://cds.cern.ch/record/2157687.

[113] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, Eur. Phys. J. C **76** (2016) 666, arXiv: 1606.01813 [hep-ex].

[114] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. C **74** (2014) 3130, arXiv: 1407.3935 [hep-ex].

[115] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex].

[116] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[117] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[118] T. Barillari et al., *Local Hadronic Calibration*, tech. rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, Due to a report-number conflict with another document, the report-number ATL-LARG-PUB-2009-001-2 has been assigned.: CERN, 2008, URL: https://cds.cern.ch/record/1112035.

[119] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex].

[120] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, (2018), arXiv: 1802.08168 [hep-ex].

[121] S. van der Meer, *Calibration of the effective beam height in the ISR*, tech. rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, 1968, URL: http://cds.cern.ch/record/296752.

[122] D0 Collaboration, *Measurement of color flow in* $\mathbf{t\bar{t}}$ *events from* $\mathbf{p\bar{p}}$ *collisions at* $\mathbf{\sqrt{s} = 1.96}$ *TeV*, Phys. Rev. **D83** (2011) 092002, arXiv: 1101.0648 [hep-ex].

[123] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s} = 8$ TeV*, Eur. Phys. J. C **75** (2015) 303, arXiv: 1412.7086 [hep-ex].

[124] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 295, arXiv: 1512.05955 [hep-ex].

[125] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2017-029, 2017, URL: https://cds.cern.ch/record/2261772.

[126] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*, Wadsworth and Brooks, 1984.

[127] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 119, ISSN: 0022-0000, URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[128] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv: hep-ph/0206293 [hep-ph].

[129] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].

[130] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. C **72** (2012) 1849, arXiv: 1110.1530 [hep-ex].

[131] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv: hep-ph/0011363 [hep-ph].

[132] G. Corcella et al., *HERWIG 6.5 release note*, (2002), arXiv: hep-ph/0210213 [hep-ph].

[133] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex].

[134] ALEPH Collaboration, *A Precise measurement of Gamma ($Z \to b\bar{b}$) / Gamma ($Z \to$ hadrons)*, Phys. Lett. **B313** (1993) 535.

[135] ATLAS Collaboration, *Commissioning of the ATLAS high performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, 2011, URL: https://cds.cern.ch/record/1369219.

[136] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex].

[137] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton–proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment*, ATLAS-CONF-2010-069, 2010, URL: https://cds.cern.ch/record/1281344.

[138] 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.

[139] W. S. McCulloch and W. Pitts,
A logical calculus of the ideas immanent in nervous activity,
The bulletin of mathematical biophysics **5** (1943) 115, ISSN: 1522-9602,
URL: https://doi.org/10.1007/BF02478259.

[140] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046, 2014,
URL: https://cds.cern.ch/record/1741020.

[141] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*,
ATLAS-CONF-2014-004, 2014, URL: https://cds.cern.ch/record/1664335.

[142] 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.

[143] ATLAS Collaboration,
A neural network clustering algorithm for the ATLAS silicon pixel detector,
JINST **9** (2014) P09009, arXiv: 1406.7690 [hep-ex].

[144] ATLAS Collaboration,
The Optimization of ATLAS Track Reconstruction in Dense Environments,
ATL-PHYS-PUB-2015-006, 2015, URL: https://cds.cern.ch/record/2002609.

[145] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*,
Springer, 2012.

[146] 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.

[147] 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.

[148] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*,
Book in preparation for MIT Press, MIT Press, 2016,
URL: http://www.deeplearningbook.org.

[149] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521** (2015),
URL: https://doi.org/10.1038/nature14539.

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

[151] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006,
2018, URL: https://cds.cern.ch/record/2314418.

[152] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001,
2018, URL: https://cds.cern.ch/record/2306649.

[153] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group,
Precision electroweak measurements on the Z resonance,
Phys. Rept. **427** (2006) 257, arXiv: hep-ex/0509008 [hep-ex].

[154] ATLAS Collaboration, *Measurement of τ polarisation in $Z/\gamma^* \to \tau\tau$ decays in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*,
Eur. Phys. J. C **78** (2018) 163, arXiv: 1709.03490 [hep-ex].

[155] S. Jadach, J. H. Kuhn, and Z. Was, *TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons*,
Comput. Phys. Commun. **64** (1990) 275.

[156] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was, and Z. Was,
Universal interface of TAUOLA: Technical and physics documentation,
Comput. Phys. Commun. **183** (2012) 821, arXiv: 1002.0543 [hep-ph].

[157] ATLAS Collaboration, *Measurement of τ polarization in $W \to \tau\nu$ decays with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **72** (2012) 2062, arXiv: 1204.6720 [hep-ex].

[158] Z. Czyczula, T. Przedzinski, and Z. Was,
TauSpinner Program for Studies on Spin Effect in tau Production at the LHC,
Eur. Phys. J. C **72** (2012) 1988, arXiv: 1201.0117 [hep-ph].

[159] A. Kaczmarska, J. Piatlicki, T. Przedziński, E. Richter-Wąs, and Z. Wąs,
Application of TauSpinner for Studies on tau-Lepton Polarization and Spin Correlations in Z, W and H Decays at the LHC,
Acta Phys. Polon. B **45** (2014) 1921, arXiv: 1402.2068 [hep-ph].

[160] S. Banerjee, J. Kalinowski, W. Kotlarski, T. Przedzinski, and Z. Was,
Ascertaining the spin for new resonances decaying into tau+ tau- at Hadron Colliders, Eur. Phys. J. **C73** (2013) 2313, arXiv: 1212.2873 [hep-ph].

[161] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*,
Comput. Phys. Commun. **77** (1993) 219.

[162] V. Cherepanov, A. Stahl, and W. Lohmann,
"Measurement of the polarization of tau–leptons produced in Z decays at CMS and determination of the effective weak mixing angle", 2016,
URL: https://cds.cern.ch/record/2206964.

[163] ATLAS Collaboration, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 658,
arXiv: 1602.04516 [hep-ex].

[164] ATLAS Collaboration,
Cross-section measurements of the Higgs boson decaying into a pair of tau-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,
Submitted to Phys. Rev. D (), arXiv: 1811.08856 [hep-ex].

[165] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (21 1967) 1264,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[166] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer,
Gluon-induced W-boson pair production at the LHC, JHEP **12** (2006) 046,
arXiv: hep-ph/0611170 [hep-ph].

[167] R. Ellis, I. Hinchliffe, M. Soldate, and J. V. D. Bij,
Higgs decay to $\tau^+\tau^-$: A possible signature of intermediate mass Higgs bosons at high energy hadron colliders, Nuclear Physics B **297** (1988) 221, ISSN: 0550-3213,
URL:
http://www.sciencedirect.com/science/article/pii/0550321388900193.

[168] A. Elagin, P. Murat, A. Pranko, and A. Safonov,
A New Mass Reconstruction Technique for Resonances Decaying to di-tau,
Nucl. Instrum. Meth. **A654** (2011) 481, arXiv: 1012.4686 [hep-ex].

[169] D. Atwood and A. Soni, *Analysis for magnetic moment and electric dipole moment form factors of the top quark via $e^+e^- \to t\bar{t}$*, Phys. Rev. D **45** (7 1992) 2405,
URL: https://link.aps.org/doi/10.1103/PhysRevD.45.2405.

[170] M. Davier, L. Duflot, F. L. Diberder, and A. Rougé,
The optimal method for the measurement of tau polarization,
Physics Letters B **306** (1993) 411, ISSN: 0370-2693, URL:
http://www.sciencedirect.com/science/article/pii/037026939390101M.

[171] M. Diehl and O. Nachtmann, *Optimal observables for the measurement of three gauge boson couplings in $e + e- \to W + W-$*,
Zeitschrift für Physik C Particles and Fields **62** (1994) 397, ISSN: 1431-5858,
URL: https://doi.org/10.1007/BF01555899.

[172] M. Ciccolini, A. Denner, and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC*,
Phys. Rev. Lett. **99** (2007) 161803, arXiv: 0707.0381 [hep-ph].

[173] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*,
Phys. Rev. **D77** (2008) 013002, arXiv: 0710.4749 [hep-ph].

[174] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck,
HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders, Comput. Phys. Commun. **195** (2015) 161,
arXiv: 1412.5390 [hep-ph].

[175] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al.,
The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,
JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[176] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 476,
arXiv: 1506.05669 [hep-ex], Erratum: Eur. Phys. J. C **76** (2016) 152.

[177] P. Nason,
A New method for combining NLO QCD with shower Monte Carlo algorithms,
JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[178] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[179] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

[180] CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons*, Phys. Lett. B **779** (2018) 283, arXiv: 1708.00373 [hep-ex].

[181] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56, arXiv: hep-ph/9704448 [hep-ph].

[182] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, and M. Duhrssen, *Measuring the Higgs Sector*, JHEP **08** (2009) 009, arXiv: 0904.3866 [hep-ph].

[183] CDF and D0 Collaborations, *Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron*, Phys. Rev. Lett. **109** (2012) 071804, arXiv: 1207.6436 [hep-ex].

[184] CMS Collaboration, *Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks*, Phys. Rev. D **89** (2014) 012003, arXiv: 1310.3687 [hep-ex].

[185] ATLAS Collaboration, *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, JHEP **01** (2015) 069, arXiv: 1409.6212 [hep-ex].

[186] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, JHEP **12** (2017) 024, arXiv: 1708.03299 [hep-ex].

[187] CMS Collaboration, *Evidence for the Higgs boson decay to a bottom quark–antiquark pair*, Phys. Lett. B **780** (2018) 501, arXiv: 1709.07497 [hep-ex].

[188] CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$*, Phys. Rev. D **92** (2015) 032008, arXiv: 1506.01010 [hep-ex].

[189] ATLAS Collaboration, *Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Phys. Rev. D **98** (2018) 052003, arXiv: 1807.08639 [hep-ex].

[190] ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **97** (2018) 072016, arXiv: 1712.08895 [hep-ex].

[191] CMS Collaboration, *Search for $t\bar{t}H$ production in the all-jet final state in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **06** (2018) 101, arXiv: 1803.06986 [hep-ex].

[192] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001) 152, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, ISSN: 0168-9002, URL: http://www.sciencedirect.com/science/article/pii/S0168900201000894.

[193] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, arXiv: 0802.2470 [hep-ph].

[194] 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.

[195] ATLAS Collaboration, *Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment*, Phys. Rev. Lett. **120** (2018) 211802, arXiv: 1802.04329 [hep-ex].

[196] ATLAS Collaboration, *Search for the Dimuon Decay of the Higgs Boson in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **119** (2017) 051802, arXiv: 1705.04582 [hep-ex].

[197] CMS Collaboration, *Search for the Higgs boson decaying to two muons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. Lett. (2018), arXiv: 1807.06325 [hep-ex].

[198] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv: 1106.0522 [hep-ph].

[199] K. P. F.R.S., *LIII. On lines and planes of closest fit to systems of points in space*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **2** (1901) 559, eprint: https://doi.org/10.1080/14786440109462720, URL: https://doi.org/10.1080/14786440109462720.

[200] I. Jolliffe, *Principal Component Analysis*, Springer, 2002, URL: https://www.springer.com/us/book/9780387954424.

[201] D. J. Hand, *Statistical Concepts: A Second Course, Fourth Edition by Richard G. Lomax, Debbie L. Hahs-Vaughn*, International Statistical Review **80** (2012) 491, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1751-5823.2012.00196_22.x, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-5823.2012.00196_22.x.

[202] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].

[203] A. Djouadi et al., *The post-Higgs MSSM scenario: Habemus MSSM?*, Eur. Phys. J. **C73** (2013) 2650, arXiv: 1307.5205 [hep-ph].

[204] M. Carena, S. Heinemeyer, O. Stal, C. E. M. Wagner, and G. Weiglein, *MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle*, Eur. Phys. J. **C73** (2013) 2552, arXiv: 1302.7033 [hep-ph].

[205] CMS Collaboration, *Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **09** (2018) 007, arXiv: 1803.06553 [hep-ex].

[206] D. M. Asner et al., "ILC Higgs White Paper", *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013, arXiv: 1310.0763 [hep-ph], URL: http://www.slac.stanford.edu/econf/C1307292/docs/submittedArxivFiles/1310.0763.pdf.

[207] CMS Collaboration, *Search for beyond the standard model Higgs bosons decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 113, arXiv: 1805.12191 [hep-ex].

[208] ATLAS Collaboration, *Search for Minimal Supersymmetric Standard Model Higgs bosons H/A and for a Z' boson in the $\tau\tau$ final state produced in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, Eur. Phys. J. C **76** (2016) 585, arXiv: 1608.00890 [hep-ex].

[209] ATLAS Collaboration, *Constraints on new phenomena via Higgs boson couplings and invisible decays with the ATLAS detector*, JHEP **11** (2015) 206, arXiv: 1509.00672 [hep-ex].

[210] L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag, 1986, URL: http://cg.scs.carleton.ca/~luc/rnbookindex.html.

[211] *Search for charged Higgs bosons with the $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, tech. rep. CMS-PAS-HIG-18-014, CERN, 2018, URL: https://cds.cern.ch/record/2640359.

[212] ATLAS Collaboration, *Search for charged Higgs bosons decaying into top and bottom quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP (2018), arXiv: 1808.03599 [hep-ex].

[213] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, ATLAS-CONF-2014-058, 2014, URL: https://cds.cern.ch/record/1951336.