Reconstruction and identification of boosted di-τ systems in a search for Higgs boson pairs using 13 TeV proton-proton collision data in ATLAS

The ATLAS Collaboration

DOI
10.1007/JHEP11(2020)163

Publication date
2020

Document Version
Final published version

Published in
Journal of High Energy Physics

License
CC BY

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Reconstruction and identification of boosted di-\(\tau\) systems in a search for Higgs boson pairs using 13 TeV proton-proton collision data in ATLAS

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: In this paper, a new technique for reconstructing and identifying hadronically decaying \(\tau^+\tau^-\) pairs with a large Lorentz boost, referred to as the di-\(\tau\) tagger, is developed and used for the first time in the ATLAS experiment at the Large Hadron Collider. A benchmark di-\(\tau\) tagging selection is employed in the search for resonant Higgs boson pair production, where one Higgs boson decays into a boosted \(b\bar{b}\) pair and the other into a boosted \(\tau^+\tau^-\) pair, with two hadronically decaying \(\tau\)-leptons in the final state. Using 139 fb\(^{-1}\) of proton-proton collision data recorded at a centre-of-mass energy of 13 TeV, the efficiency of the di-\(\tau\) tagger is determined and the background with quark- or gluon-initiated jets misidentified as di-\(\tau\) objects is estimated. The search for a heavy, narrow, scalar resonance produced via gluon-gluon fusion and decaying into two Higgs bosons is carried out in the mass range 1–3 TeV using the same dataset. No deviations from the Standard Model predictions are observed, and 95% confidence-level exclusion limits are set on this model.

KEYWORDS: Beyond Standard Model, Hadron-Hadron scattering (experiments), Higgs physics, Tau Physics

ArXiv ePrint: 2007.14811
1 Introduction

The discovery of the Higgs boson (H) by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) in 2012 [1, 2] opens new ways of probing physics beyond the Standard Model (SM), since the Higgs boson may itself appear as one of the intermediate states in the decay of new resonances. Various final states have been used by ATLAS and CMS in searches for both resonant and non-resonant HH production [3–8]. In the high-mass regime, for resonance masses typically above 1 TeV, the Higgs bosons may be produced with large momenta, causing their decay products to be collimated. The standard reconstruction techniques become inefficient in this regime. Therefore, a new technique, referred to as the di-τ tagger, is developed to reconstruct and identify boosted hadronically decaying τ+τ− pairs. For the identification, a multivariate algorithm is trained to distinguish between τ+τ− pairs and the multi-jet background from quark- or gluon-initiated jets by exploiting the calorimetric shower shapes and tracking information. A similar algorithm was implemented by the CMS Collaboration in ref. [9].
An application of the di-τ tagger is carried out in a search for a narrow spin-0 resonance in the mass range 1–3 TeV, which is produced via gluon-gluon fusion and decays into a pair of Higgs bosons ($X \rightarrow HH$), as predicted by models with an extended Higgs sector, such as two-Higgs-doublet models [10, 11]. This search considers the final state where one Higgs boson decays into a $b\bar{b}$ pair and the other one into a $\tau^+\tau^-$ pair, where both $\tau$-leptons decay hadronically.1 A dedicated benchmark of the di-τ tagger with an identification efficiency of 60% is designed for this analysis. Using 139 fb$^{-1}$ of proton-proton (pp) collision data at a centre-of-mass energy $\sqrt{s} = 13$ TeV recorded by the ATLAS experiment in 2015–2018, various orthogonal event categories are defined in order to correct the efficiency of the di-τ tagger for the benchmark selection, to perform and validate the multi-jet background estimate, and to search for resonantly produced Higgs boson pairs.

This paper is organised as follows. After a brief description of the ATLAS detector in section 2, the samples of data and simulated events used in this study are described in section 3. The procedures used to reconstruct and identify physics objects such as electrons, muons, jets and missing transverse momentum in the detector are described in section 4. Section 5 presents the reconstruction and identification of boosted hadronically decaying $\tau^+\tau^-$ pairs. In section 6, general event selections and categorisations are summarised, while section 7 focuses on the data-driven estimation of the multi-jet background with quark- or gluon-initiated jets misidentified as boosted hadronically decaying $\tau^+\tau^-$ pairs. The data-driven correction of the di-τ tagger efficiency is discussed in section 8 and the search for $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ is presented in section 9, including the statistical analysis used to set 95% confidence-level (CL) limits on the production cross-section for resonant $HH$ production. Finally, a summary is given in section 10.

2 ATLAS detector

The ATLAS detector [12] at the LHC is a multipurpose particle detector with a forward-backward symmetric cylindrical geometry and nearly 4π coverage in solid angle.2 It consists of an inner tracking system surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer.

The inner detector covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. For the $\sqrt{s} = 13$ TeV run, a fourth layer of the pixel detector, the insertable $B$-layer [13, 14], was installed close to the beam pipe at an average radius of 33.2 mm, providing an additional position measurement with 8 $\mu$m resolution in the $(x,y)$ plane and 40 $\mu$m along $z$.

Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic energy measurements with high granularity in the region $|\eta| < 3.2$. In the central part, $|\eta| < 2.5$, the calorimeter is divided into three layers, one of them segmented in thin $\eta$ strips for optimal

---

1SM branching fractions of the Higgs boson are assumed throughout the paper.

2ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upwards. Cylindrical coordinates $(r,\phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. 

---
\(\gamma/\pi_0\) separation, completed by a presampler layer for |\(\eta| < 1.8\). A hadronic steel/scintillator-tile calorimeter covers the central pseudorapidity range (|\(\eta| < 1.7\)). The endcap and forward regions are instrumented with LAr calorimeters for both the electromagnetic and hadronic energy measurements up to |\(\eta| = 4.9\). The granularity of the calorimeter system in terms of \(\Delta \eta \times \Delta \phi\) is typically 0.025 \(\times \pi/128\) in the barrel of the electromagnetic calorimeter and 0.1 \(\times \pi/32\) in the hadronic calorimeter, with variations in segmentation with |\(\eta| and the layer [15].

The muon spectrometer surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. The muon spectrometer includes a system of precision tracking chambers and fast detectors for triggering.

A two-level trigger system [16] is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 100 kHz. This is followed by a software-based trigger that reduces the accepted event rate to 1 kHz on average.

### 3 Data and simulated events

The studies presented in this paper are performed using a sample of pp collision data recorded at a centre-of-mass energy \(\sqrt{s} = 13\) TeV between 2015 and 2018, during stable beam conditions and when all detector components relevant to the analysis were operating nominally [17]. This corresponds to an integrated luminosity of 139 fb\(^{-1}\). Samples of Monte Carlo (MC) simulated events are used to train and calibrate the di-\(\tau\) tagger, as well as to model the signal and some SM background processes in the search for resonant Higgs boson pair production.

The signal, i.e. the production of a heavy spin-0 resonance via gluon-gluon fusion and its decay into a pair of Higgs bosons, \(X \to HH\), was simulated for nine values of the resonance mass, \(m_X\), between 1 and 3 TeV, using MADGRAPH5_aMC@NLO v2.6.1 [18] at leading-order (LO) accuracy in quantum chromodynamics (QCD) with the NNPDF2.3LO [19] set of parton distribution functions (PDFs). The event generator was interfaced with HERWIG v7.1.3 [20, 21] to model the parton shower, hadronisation and underlying event, using the default set of tuned parameters (tune) and the MMHT2014LO [22] PDF set. In the nine signal samples, a narrow-width approximation was used for the resonance \(X\), i.e. its natural width was set to a value that remains much smaller than the experimental mass resolution. In addition, the Higgs boson mass was set to 125 GeV and the SM branching fractions were used for the decays \(H \to b\bar{b}\) and \(H \to \tau^+\tau^-\). In order to develop the identification algorithm of the boosted hadronically decaying \(\tau^+\tau^-\) pairs, another set of \(HH\) samples was produced, based on a narrow-width spin-2 Kaluza-Klein graviton, as predicted in the Randall-Sundrum model of warped extra dimensions [23], \(G \to HH \to (\tau^+\tau^-)(\tau^+\tau^-)\). Such events were generated with MADGRAPH5_aMC@NLO v2.3.3 at LO accuracy in QCD with the NNPDF2.3LO set of PDFs, interfaced with PYTHIA v8.212 [24] using the A14 [25] tune. Five samples, with graviton masses of 1.5, 1.75, 2, 2.25 and 2.5 TeV, were generated. Only hadronically decaying \(\tau\)-leptons were considered in all \(HH\) signal samples.
The production of $W$ and $Z$ bosons in association with jets ($V$+jets) was simulated with Sherpa v2.2.1 [26] using matrix elements at next-to-leading-order (NLO) accuracy in QCD for up to two jets and at LO accuracy for up to four jets, calculated with the Comix [27] and OpenLoops [28] libraries. They were matched with the Sherpa parton shower [29] using the MEPS@NLO prescription [30, 31]. The tune developed by the Sherpa authors and the NNPDF3.0NNLO PDF set were used. The $V$+jets samples were normalised to a next-to-next-to-leading-order (NNLO) prediction [32]. In the $Z$+jets events, jets are labelled according to the generated hadrons with $p_T > 5$ GeV found within a cone of size $\Delta R = 0.4$ around the jet axis. If a $b$-hadron is found, the jet is labelled as a $b$-jet. If no $b$-hadron is found but there is a $c$-hadron instead, the jet is labelled as a $c$-jet. If neither a $b$-hadron nor a $c$-hadron is found, the jet is labelled as a light ($l$) jet. Simulated $Z$+jets events are then categorised according to the labels of the two jets that are used to reconstruct the $H \to b\bar{b}$ candidate. The combination of $Z+b\bar{b}$, $Z+bc$, $Z+bl$ and $Z+cc$ events is referred to as $Z+hf$ (denoting heavy-flavour jets) in the following, whereas other events belong to the $Z+lf$ (denoting light-flavour jets) category. This categorisation is not performed for the $W$+jets process because its contribution is small.

The Powheg-Box v2 generator [33–35] was used to generate the $WW$, $WZ$ and $ZZ$ (diboson) processes [36] at NLO accuracy in QCD. The effect of singly resonant amplitudes, as well as interference effects due to $Z/\gamma^*$ and identical leptons in the final state, were included where appropriate (interference effects between $WW$ and $ZZ$ for same-flavour charged leptons and neutrinos were ignored). Events were interfaced with Pythia v8.186 [37] for the modelling of the parton shower, hadronisation and underlying event, with parameters set according to the AZNLO [38] tune. The CT10 [39] PDF set was used for the hard-scattering processes, whereas the CTEQ6L1 [40] PDF set was used for the parton shower.

The production of a single 125 GeV Higgs boson in association with a $Z$ boson was simulated up to NLO accuracy in QCD using Powheg-Box v2 [41–43], with the NNPDF3.0NLO PDF set and subsequently reweighted to the PDF4LHC15NLO [44] PDF set. The simulation was interfaced with Pythia v8.212, using the AZNLO tune and the CTEQ6L1 PDF set. The $gg \to ZH$ and $qq \to ZH$ samples were normalised to cross-sections calculated at, respectively, NLO accuracy in QCD including soft-gluon resummation up to next-to-leading logarithms [45–47] and NNLO accuracy in QCD with NLO electroweak corrections [48–55]. Other single-Higgs-boson production modes were found to contribute negligibly.

Single-top-quark processes (split into $s$-channel, $t$-channel and $tW$ contributions) and $t\bar{t}$ events were simulated using Powheg-Box v2 [56–58] at NLO accuracy in QCD with the NNPDF3.0NLO PDF set. All events were interfaced with Pythia v8.230 using the A14 tune and the NNPDF2.3LO PDF set. For the $tW$ process, the diagram removal scheme [59] was employed in order to handle the interference with $t\bar{t}$ production. The $t\bar{t}$ sample was normalised to the cross-section prediction at NNLO accuracy in QCD including the resummation of next-to-next-to-leading logarithmic soft-gluon terms calculated using Top++2.0 [60–66]. For the single-top-quark processes, the cross-sections of the $s$- and $t$-channels were corrected to the theory prediction at NLO accuracy in QCD calculated with Hathor v2.1 [67, 68], while the cross-section used for the $tW$ sample was based on the approximate NNLO calculation [69, 70].
Except when using Sherpa, $b$- and $c$-hadron decays were performed with EvtGen v1.2.0 or v1.6.0 [71], while the decays of $\tau$-leptons were handled internally by all event generators. The effect of multiple interactions in the same and neighbouring bunch crossings (pile-up) was modelled by overlaying the original hard-scattering event with simulated inelastic $pp$ events generated with Pythia v8.186 using the NNPDF2.3LO PDF set and the A3 [72] tune. The MC samples were processed with a simulation [73] of the detector response based on Geant4 [74] and events were then reconstructed with the same software as the data.

4 Object reconstruction

The following procedures are used to reconstruct and identify objects, such as electrons, muons, jets and missing transverse momentum, in the ATLAS experiment.

In general, jets refer to the hadronic objects reconstructed using the anti-$k_T$ algorithm with a radius parameter $R = 0.4$ [75, 76], starting from topological clusters of energy deposits in the calorimeter. When used in the following, jets are required to have $p_T > 20\text{ GeV}$ and $|\eta| < 4.5$. The reconstruction of electrons is based on matching inner-detector tracks to energy clusters in the electromagnetic calorimeter. Electrons are required to have $p_T > 7\text{ GeV}$ and $|\eta| < 2.47$, excluding the barrel-endcap transition region of the calorimeter ($1.37 < |\eta| < 1.52$). They are then identified using the ‘loose’ operating point provided by a likelihood-based algorithm [77]. The reconstruction of muons relies on matching tracks in the inner detector and the muon spectrometer. Muons are required to have $p_T > 7\text{ GeV}$ and $|\eta| < 2.5$, as well as to satisfy the ‘loose’ identification criteria and ‘FixedCutLoose’ isolation working point defined in ref. [78].

In order to avoid double-counting of objects that overlap geometrically, an electron is removed if it shares an inner-detector track with a muon. Then, anti-$k_T$ jets with $R = 0.4$ are discarded if they meet one of the following two conditions: (i) $\Delta R(\text{jet}, e) < 0.2$; (ii) the jet has less than three associated tracks and either a muon inner-detector track is associated with the jet or $\Delta R(\text{jet}, \mu) < 0.2$. Finally, an electron or muon is discarded if found within a distance $\Delta R = \min(0.4, 0.04 + 10\text{ GeV}/p_T^{e/\mu})$ from any remaining jet.

The missing transverse momentum, the magnitude of which is denoted by $E_T^{\text{miss}}$ in the following, is defined as the negative vector sum of the transverse momenta of all fully reconstructed and calibrated objects, after an overlap removal procedure that is distinct from that used for the electron/muon/jet disambiguation above [79, 80]. The missing transverse momentum also includes a soft term, calculated using the inner-detector tracks that originate from the primary vertex (defined as that having the largest sum of squared track-$p_T$) but are not associated with reconstructed objects.

In order to capture the decay products of boosted particles, such as Higgs bosons, another type of jets, called large-radius jets, is employed. These are also formed using the anti-$k_T$ algorithm, but with a radius parameter $R = 1.0$ and are built from topological clusters of energy deposits calibrated using the local hadronic cell weighting scheme [15]. In the following, the large-radius jet matched to the boosted $H \rightarrow b\bar{b}$ candidate is referred to as a ‘large-$R$ jet’ while that of the boosted hadronically decaying $\tau^+\tau^-$ pair is referred to as a ‘di-$\tau$ object’.
While the reconstruction and identification of boosted $H \rightarrow \tau^+ \tau^-$ decays employ new techniques described in section 5, a standard procedure is used for boosted $H \rightarrow b\bar{b}$ decays. Large-$R$ jets are trimmed [81] to remove the effects of pile-up and the underlying event. Trimming proceeds by reclustering the original constituents of a large-$R$ jet into a collection of $R = 0.2$ sub-jets using the $k_t$ algorithm [82, 83] and removing any sub-jets with $p_T^{\text{sub-jet}}/p_T^{\text{jet}} < 0.05$, where $p_T^{\text{sub-jet}}$ is the transverse momentum of the sub-jet under consideration and $p_T^{\text{jet}}$ that of the original (untrimmed) large-$R$ jet. The energy and mass scales of the trimmed jets are then calibrated using $p_T$- and $\eta$-dependent calibration factors derived from simulation [84]. After trimming, the large-$R$ jet is required to have $p_T > 300$ GeV and its mass is calculated using the combined mass technique with tracking and calorimeter information as input [85].

In order to identify $b$-hadrons within a large-$R$ jet, variable-radius track-jets [86] are reconstructed using the anti-$k_t$ algorithm from inner-detector tracks with a jet-$p_T$ dependent radius parameter $R(p_T) = \rho/p_T$, where $\rho$ determines how fast the effective size of the jet decreases with its transverse momentum. The lower ($R_{\text{min}}$) and upper ($R_{\text{max}}$) cut-offs prevent the jet from becoming too large at low $p_T$ and from shrinking below the detector resolution at high $p_T$, respectively. In this paper, $\rho = 30$ GeV, $R_{\text{min}} = 0.02$ and $R_{\text{max}} = 0.4$ are used [87]. These track-jets are then matched to an untrimmed large-$R$ jet by using the ghost-association method [88, 89]. Only track-jets with $p_T > 10$ GeV, $|\eta| < 2.5$ and at least two tracks [87] are considered in the following. Events with two collinear track-jets $a$ and $b$ that fulfil $\Delta R(\text{jet}_a, \text{jet}_b) < \min(R_{\text{jet}_a}, R_{\text{jet}_b})$ are removed.

The flavour of track-jets is determined using a multivariate approach based on the properties and vertex information of the associated tracks [90, 91]. Various $b$-jet identification algorithms are used to exploit impact-parameter information, secondary-vertex information, and $b$- to $c$-hadron decay chain information. The MV2c10 algorithm [92] then combines information from the various upstream algorithms in a boosted decision tree (BDT) that is trained to discriminate $b$-jets from a background sample made of 93% light-flavour jets and 7% $c$-jets. In order to label a track-jet as $b$-tagged, a requirement is placed on the output score of the MV2c10 discriminant. This requirement has an average efficiency of 70% for $b$-jets in simulated $t\bar{t}$ events (decreasing to about 60% for a $b$-jet $p_T$ above 500 GeV) with rejection factors of 8.9, 36 and 300 for jets initiated by $c$-quarks, hadronically decaying $\tau$-leptons and light-flavour quarks, respectively [93]. The number of $b$-tagged track-jets in the large-$R$ jet is used to define various event categories, as described in section 6.

5 Reconstruction and identification of boosted hadronically decaying $\tau^+\tau^-$ pairs

Hadronically decaying $\tau$-leptons produce a neutrino and visible decay products, typically one or three charged pions and up to several neutral pions, which are reconstructed and identified as $\tau_{\text{had-vis}}$ objects. In the standard procedure [94], a $\tau_{\text{had-vis}}$ object is seeded by a jet with $p_T > 10$ GeV and $|\eta| < 2.5$, formed using the anti-$k_t$ algorithm with $R = 0.4$. A multivariate identification stage, using calorimetric shower shapes and tracking information as input variables, is employed to discriminate $\tau_{\text{had-vis}}$ candidates with one or
Figure 1. Schematic representation of a di-τ object: the large blue cone is the di-τ seeding jet while the two smaller yellow cones are the two leading sub-jets. Tracks found within a cone of size $\Delta R = 0.2$ around the sub-jet axis are matched to charged pions produced in hadronic decays of τ-leptons, while other tracks found in the isolation region are labelled as ‘iso-tracks’. The closest distance $d_0$ in the transverse plane between the primary vertex and the leading track matched to a sub-jet is also shown for illustration.

three associated tracks from quark- or gluon-initiated jets. However, in the search for high-mass $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ presented in this paper, more than 50% of the $\tau^+\tau^-$ pairs have a separation $\Delta R < 0.4$ when $m_X \geq 2$ TeV, hence they would fail the standard reconstruction procedure. For such events, a new method for reconstructing boosted hadronically decaying $\tau^+\tau^-$ pairs, referred to as the di-τ tagger, is employed.

Boosted di-τ objects are seeded by untrimmed large-radius jets that must have $p_T > 300$ GeV. Their constituents are reclustered into anti-$k_t$ sub-jets with $R = 0.2$. The original di-τ seeding jet must include at least two such sub-jets and, after ordering in $p_T$, the two leading sub-jets are used to construct the di-τ system. Tracks are geometrically matched to a sub-jet if they are within a cone of size $\Delta R = 0.2$ around its axis, and they are labelled as ‘τ tracks’. Other tracks found in the isolation region (i.e. the area of the di-τ seeding large-radius jet excluding the di-τ sub-jets) are labelled as ‘iso-tracks’. A schematic representation of a di-τ object is shown in figure 1. The track selection criteria, as well as the track-vertex matching, are the same as those used for the standard $\tau_{had-vis}$ objects [94]. In the following, the two leading sub-jets used to compute the four-momentum of the di-τ system must have $p_T > 10$ GeV and at least one associated track.

At this stage, the di-τ reconstruction efficiency is defined as the fraction of events in which a boosted di-τ candidate is reconstructed and each of the two leading sub-jets geometrically matches a generated hadronically decaying τ-lepton for which the $p_T$ of the visible products (neutral and charged hadrons) exceeds 10 GeV. Figures 2(a) and 2(b) show the di-τ reconstruction efficiency as a function of, respectively, the distance $\Delta R(\tau_{1,vis}, \tau_{2,vis})$ between the visible products of the two hadronically decaying τ-leptons and the visible $p_T$ of the di-τ system, both computed at generator level. For this purpose, a sample of simulated $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ events is used, in which the resonance mass is set to 2 TeV and both τ-leptons decay hadronically. In figure 2, the reconstruction efficiency is also shown for
standard (resolved) \( \tau_{\text{had-vis}} \) objects. In this case, the efficiency is defined as the fraction of events with at least two reconstructed \( \tau_{\text{had-vis}} \) candidates, with at least one associated track each, that geometrically match a generated hadronically decaying \( \tau \)-lepton for which the \( p_T \) of the visible products exceeds 10 GeV. This comparison shows that the boosted di-\( \tau \) object reconstruction method is necessary at high transverse momenta and low \( \Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}}) \). In particular, figure 2(a) shows that the reconstruction efficiencies decrease sharply when the visible products of the two hadronically decaying \( \tau \)-leptons are so close that they merge into one jet. With the resolved \( \tau_{\text{had-vis}} \) reconstruction this happens when \( \Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}}) < 0.4 \), while the boosted di-\( \tau \) reconstruction extends the sensitivity down to \( \Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}}) = 0.2 \) by resolving the smaller sub-jets. In addition, as the distance \( \Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}}) \) increases, the reconstruction efficiency of both boosted di-\( \tau \) and resolved \( \tau_{\text{had-vis}} \) objects decreases slowly, because the sub-leading generated \( \tau \)-lepton is found to become softer, and hence less likely to exceed the 10 GeV \( p_T \) threshold imposed on its visible products. Also, in contrast to the reconstruction efficiency of resolved \( \tau_{\text{had-vis}} \) objects, which is based on at least two candidates, the di-\( \tau \) reconstruction method loses some efficiency due to the fact that only the two leading sub-jets are considered. Finally, as shown in figure 2(b), the reconstruction efficiency reaches a plateau when the \( p_T \) of the di-\( \tau \) system exceeds 300 GeV, while the location of the turn-on is set by the \( p_T \) cut on the seeding jet.

As in the case of standard \( \tau_{\text{had-vis}} \) objects, a separate identification stage using multivariate techniques is employed to reduce the background from quark- and gluon-initiated jets. For this purpose, a BDT discriminant is built using information about the clusters in the calorimeter, tracks and vertices. Multi-jet events with quark- or gluon-initiated jets misidentified as di-\( \tau \) objects are expected to have lower-\( p_T \) sub-jets, with a larger fraction of energy in the isolation region. Also, multi-jet events typically have fewer collimated
Table 1. Discriminating variables used in the di-τ identification BDT, aimed at rejecting the background from quark- and gluon-initiated jets. Here, LRJ refers to the seeding large-radius jet of the di-τ object, s_j1 and s_j2 stand for the first and second sub-jets ordered in p_T, respectively, and tracks refer to those matched to a sub-jet (τ tracks), unless specified otherwise.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>E^{s_j1}<em>{ΔR&lt;0.1}/E^{s_j1}</em>{ΔR&lt;0.2} and E^{s_j2}<em>{ΔR&lt;0.1}/E^{s_j2}</em>{ΔR&lt;0.2}</td>
<td>Ratios of the energy deposited in the core to that in the full cone, for the sub-jets s_j1 and s_j2, respectively</td>
</tr>
<tr>
<td>p_T^{s_j2}/p_T^{LRJ} and (p_T^{s_j1} + p_T^{s_j2})/p_T^{LRJ}</td>
<td>Ratio of the p_T of s_j2 to the di-τ seeding large-radius jet p_T and ratio of the scalar p_T sum of the two leading sub-jets to the di-τ seeding large-radius jet p_T, respectively</td>
</tr>
<tr>
<td>log(∑p_T^{iso-tracks}/p_T^{LRJ})</td>
<td>Logarithm of the ratio of the scalar p_T sum of the iso-tracks to the di-τ seeding large-radius jet p_T</td>
</tr>
<tr>
<td>ΔR_{max}(track,s_j1) and ΔR_{max}(track,s_j2)</td>
<td>Largest separation of a track from its associated sub-jet axis, for the sub-jets s_j1 and s_j2, respectively</td>
</tr>
<tr>
<td>∑</td>
<td>p_T^{track}ΔR(track,s_j2)</td>
</tr>
<tr>
<td>∑</td>
<td>p_T^{iso-track}ΔR(iso-track,s_j)</td>
</tr>
<tr>
<td>log(m_{ΔR&lt;0.1}^{tracks,s_j1}) and log(m_{ΔR&lt;0.1}^{tracks,s_j2})</td>
<td>Logarithms of the invariant mass of the tracks in the core of s_j1 and s_j2, respectively</td>
</tr>
<tr>
<td>log(m_{ΔR&lt;0.2}^{tracks,s_j1}) and log(m_{ΔR&lt;0.2}^{tracks,s_j2})</td>
<td>Logarithms of the invariant mass of the tracks with ΔR &lt; 0.2 from the axis of s_j1 and s_j2, respectively</td>
</tr>
<tr>
<td>log(</td>
<td>d_0^{s_j1}</td>
</tr>
<tr>
<td>n_{tracks}^{s_j1} and n_{sub-jets}^{tracks}</td>
<td>Number of tracks matched to s_j1 and to all sub-jets, respectively</td>
</tr>
</tbody>
</table>

tracks and a higher track multiplicity with, accordingly, a smaller fraction of the transverse momentum carried by each track. The variables defined in table 1 are found to provide good discrimination and are used in the BDT training.

The training of the BDT is performed using the adaptive boosting (AdaBoost) algorithm [95], with a boosting parameter of 0.5, in the TMVA toolkit [96]. The di-τ objects in simulated events with a spin-2 graviton, G → HH → (τ⁺τ⁻)(τ⁺τ⁻), form the signal training set. Five samples, with graviton masses of 1.5, 1.75, 2, 2.25 and 2.5 TeV, are combined to form the signal. The spin of the resonance is found to have no impact on the di-τ identification.

The background sample in the BDT training consists of the 3.2 fb⁻¹ of data recorded by the ATLAS experiment in 2015 using combined jet triggers with transverse energy (E_T) thresholds between 100 and 400 GeV. It is dominated by multi-jet events and the contamination from events with a true di-τ object is found to be negligible. The di-τ p_T
Figure 3. (a) Background rejection factor versus identification efficiency for the BDT used to discriminate between di-τ signatures and quark- or gluon-initiated jets, as obtained after the di-τ reconstruction step (the red cross indicates the benchmark used in the analysis, corresponding to a signal efficiency of about 60% and a background rejection factor of $10^4$), and identification efficiency of boosted di-τ objects, as a function of (b) their $p_T$ at generator level and (c) the number $\mu$ of pile-up interactions. The vertical error bars only account for statistical uncertainties.

The spectra of the signal and background events are reweighted to become flat, so that the $p_T$ dependencies of the input variables are eliminated in the BDT training. On the other hand, no reweighting regarding $\eta$ is performed, since no dependency on this variable is observed. Figure 3(a) shows the background rejection factor (defined as the inverse of the background efficiency) versus the signal efficiency, as obtained after the di-τ reconstruction step. In the following, a requirement is placed on the BDT output score to define a benchmark with a signal efficiency of about 60%, as indicated by the red cross. This working point corresponds to a background rejection factor of $10^4$. Figure 3(b) shows that the identification efficiency has little dependency on the $p_T$ of the di-τ system beyond 300 GeV. As illustrated in figure 3(c), the performance of the BDT is not affected much by the number of pile-up interactions. Hence, although the training is performed using only data collected in 2015, the discrimination between the di-τ signature and the multi-jet background is not expected to change significantly across the data-taking period between 2015 and 2018.
Figure 4. Efficiency of the track selection for di-τ sub-jets versus (a) the $p_T$ of the di-τ system at generator level and (b) the number $\mu$ of pile-up interactions, computed after the boosted di-τ reconstruction and identification steps in simulated $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ events, where the resonance mass is set to 2 TeV and both τ-leptons decay hadronically. The vertical error bars only account for statistical uncertainties.

Since the majority of hadronically decaying τ-leptons produce one or three charged pions, the number of tracks matched to each di-τ sub-jet, with a distance $\Delta R < 0.1$ from its axis, must be either one or three in the following. The efficiency of this additional requirement, computed after the di-τ reconstruction and identification steps, is shown in figures 4(a) and 4(b) as a function of the generator-level $p_T$ of the di-τ system and the number $\mu$ of pile-up interactions, respectively. For this purpose, the same sample of simulated $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ events is used. The di-τ sub-jet track selection is found to be stable not only with respect to the $p_T$ of the di-τ system and the number of pile-up interactions, but also with respect to pseudorapidity and the distance $\Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}})$ between the visible decay products of the two hadronically decaying τ-leptons. This additional requirement further improves the rejection of the background with quark- or gluon-initiated jets by a factor of about five.

In simulated $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ events, the energy of reconstructed di-τ objects is close to that of the visible decay products of the corresponding two hadronically decaying τ-leptons at generator level, as illustrated in figure 5. In addition, the energy resolution is found to improve with increasing energy. In regions of the data enriched in boosted di-τ candidates, good agreement is found between simulated and measured di-τ energy distributions; hence no additional energy calibration is applied to di-τ objects.

6 Event selection and categorisation

This section presents the common event selection and further categorisations used for the multi-jet background estimation, the data-driven correction of the efficiency of the di-τ tagger and the search for $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$. The events are selected using unprescaled triggers requiring the presence of a large-radius jet with an $E_T$ threshold between 360 and
Figure 5. Energy (a) scale and (b) resolution of the di-τ reconstruction as a function of the $p_T$ of the di-τ system at generator level, computed in simulated $X \to HH \to b\bar{b}\tau^+\tau^-$ events, where the resonance mass is set to 2 TeV and all τ-leptons decay hadronically. The requirement of $p_T > 300$ GeV on the seeding jet truncates the reconstructed energy distribution and leads to a degradation of the energy scale for the lower values of the $p_T$ of the di-τ system. The vertical error bars only account for statistical uncertainties.

460 GeV, depending on the data-taking period. Subsequently, in order to ensure a constant trigger efficiency, an offline $p_T$ cut that exceeds the trigger threshold by 40–50 GeV is applied to the leading large-radius jet, independently of whether it is compatible with a boosted $b\bar{b}$ or hadronically decaying $\tau^+\tau^-$ pair. In order to avoid contamination from non-collision backgrounds, such as those originating from calorimeter noise, beam halo and cosmic rays, events are rejected if they contain an anti-$k_t$ jet with $R = 0.4$ and $p_T > 20$ GeV that does not fulfil the loose quality criteria of ref. [97]. In addition, events are required to have a vertex with at least two associated tracks with $p_T > 500$ MeV. Finally, events are vetoed if they contain an electron or muon that meets the requirements described in section 4.

At preselection, events are required to contain at least one reconstructed di-τ object, which must meet the requirements below, in addition to those listed in section 5. If multiple di-τ objects are found, the one with the highest $p_T$ is selected.

- The number of sub-jets is at most three, and the $p_T$ threshold of the two leading sub-jets is raised to 50 GeV in order to reduce the contribution of multi-jet background events.

- The separation $\Delta R$ between the leading and sub-leading sub-jets is less than 0.8 to ensure that both are fully contained in the di-τ seeding jet.

- The charge product of the two leading sub-jets is $Q = q^{lead} \times q^{sub-lead} = \pm 1$. The charge of the sub-jet is defined as the sum of the charges of the associated tracks.

- The transverse momentum of the di-τ object, computed from the two leading sub-jets, must exceed 300 GeV (or fulfil the tighter $p_T$ requirement set by the trigger threshold if the seeding large-radius jet fires the trigger), with a pseudorapidity $|\eta| < 1.37$. 
Table 2. Definition of various categories after the event preselection (see text for details), based on the charge product $Q$ of the two leading di-$\tau$ sub-jets (OS for opposite-sign, SS for same-sign), the number $N_{b\text{-tags}}$ of $b$-tagged track-jets in the selected large-$R$ jet, the azimuthal angle $|\Delta \phi|$ between the di-$\tau$ object and the missing transverse momentum, the large-$R$ jet mass $m_J$, and the visible mass $m_{\text{vis}}^{HH}$ of the reconstructed $HH$ system.

| Name          | Usage                                      | $Q$ | $N_{b\text{-tags}}$ | $|\Delta \phi|$ | $m_J$ [GeV] | $m_{\text{vis}}^{HH}$ [GeV] |
|---------------|--------------------------------------------|-----|---------------------|-----------------|-------------|-----------------------------|
| FF SS         | Fake-factor (FF) computation               | +1  | 0                   | —               | —           | —                           |
| MJ OS 0-tag   | Closure test, multi-jet 0-$b$-tag           | —1  | 0                   | $>1$            | —           | —                           |
| MJ OS 1-tag   | Closure test, multi-jet 1-$b$-tag           | —1  | 1                   | $>1$            | 50–60 or $>160$ | $<1500$                  |
| $Z\tau\tau$ 0-tag | Di-$\tau$ tagger correction               | —1  | 0                   | $<1$            | —           | —                           |
| $Z\tau\tau$ 1-tag | Di-$\tau$ tagger closure test              | —1  | 1                   | $<1$            | 50–60 or $>160$ | $<1500$                  |
| $HH$ signal region | $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ search | —1  | 2                   | $<1$            | $>60$ and $<160$ | $[0,900,1200]$ |

or $1.52 < |\eta| < 2.0$, thereby rejecting the transition region between the barrel and endcap calorimeters.

In addition, $E_T^{\text{miss}} > 10$ GeV is required in order to define a direction for the missing transverse momentum. To select boosted $H \rightarrow b\bar{b}$ decays, large-$R$ jets with a $p_T$ exceeding 300 GeV (or the tighter trigger-dependent threshold if the seeding large-radius jet fires the trigger), $|\eta| < 2.0$, a combined mass larger than 50 GeV and a separation $\Delta R > 1.0$ from the selected di-$\tau$ object are required. If multiple large-$R$ jets fulfil these requirements, the one with the highest $p_T$ is selected.

Events are then further divided into regions with either misidentified or true di-$\tau$ objects. In the latter case, the event categories contain a significant fraction of di-$\tau$ objects from either $Z \rightarrow \tau^+\tau^-$ or $H \rightarrow \tau^+\tau^-$ decays. As shown in table 2, the categorisation is based on the charge product $Q$ of the two leading di-$\tau$ sub-jets, the number $N_{b\text{-tags}}$ of $b$-tagged track-jets in the selected large-$R$ jet, the azimuthal angle $|\Delta \phi|$ between the selected di-$\tau$ object and the missing transverse momentum, the combined mass $m_J$ of the selected large-$R$ jet, and the invariant mass of the selected di-$\tau$ object and the large-$R$ jet, $m_{\text{vis}}^{HH}$. Requirements on $m_J$ and $m_{\text{vis}}^{HH}$ are imposed in regions with one $b$-tagged large-$R$ jet in order to suppress the contamination from $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$ signal events. Finally, in order to ensure orthogonality to searches for $X \rightarrow HH \rightarrow b\bar{b}bb$ [98], events with two or more $b$-tagged large-$R$ jets are removed.

### 7 Estimation of the multi-jet background with a misidentified di-$\tau$ object

Multi-jet events with quark- or gluon-initiated jets misidentified as di-$\tau$ objects are a common background for the efficiency measurement of the di-$\tau$ tagger and in the search for $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$. This section describes the data-driven method developed to estimate this background, whereas non-multi-jet background events with a misidentified di-$\tau$ object are obtained directly from simulation.

A data sample enriched in misidentified di-$\tau$ objects is collected using the same event selection as in a region of interest (RoI), except that the selected di-$\tau$ object must fail
Figure 6. Fake factor as a function of the $p_T$ of the di-$\tau$ object, computed in the control region labelled as ‘FF SS’ in table 2, with the requirements of $Q = +1$ for the di-$\tau$ object and of zero $b$-tagged track-jets for the selected large-$R$ jet. The error bars indicate the statistical uncertainties.

(instead of pass) the BDT-based identification, but it still fulfils a very loose criterion that corresponds to a cut on the BDT output score with an efficiency above 99%. In fact, this very loose criterion is applied to all di-$\tau$ candidates in order to keep the composition of misidentified di-$\tau$ objects in terms of quarks and gluons close to that in the RoI. The contribution from non-multi-jet processes, obtained using simulation, is subtracted from the data. The remaining events, defined as $N_{\text{fail}}$, are then reweighted by a fake factor (FF) to predict the yield of events $N_{\text{mis-ID}}$ with a misidentified di-$\tau$ object in the RoI as follows:

$$N_{\text{mis-ID}} = N_{\text{fail}} \times \text{FF}.$$  

The FF is calculated in the region labelled as ‘FF SS’ in table 2, which requires $Q = +1$ for the di-$\tau$ object and is enriched in multi-jet events. It is binned in the $p_T$ of the di-$\tau$ object and defined as:

$$\text{FF} = \frac{N_{\text{pass}}}{N_{\text{fail}}},$$

where the numerator is the number of events passing the BDT-based di-$\tau$ identification requirement, while the denominator is the number of events failing this requirement but passing the very loose BDT requirement. The contributions from non-multi-jet processes in the numerator and the denominator are about 24% and 3% of the total event yields, respectively, and they are subtracted. Figure 6 shows the FF as a function of the $p_T$ of the di-$\tau$ object. The observed increase of the FF relates to a reduced rejection of the background from quark- or gluon-initiated jets at higher $p_T$ values of the di-$\tau$ object. On the other hand, the FF is inclusive in $\eta$, since no significant change in the multi-jet background modelling was found with an additional binning in the pseudorapidity of the di-$\tau$ object. The statistical uncertainty of the measured FF comes from the limited size of the samples of data and simulated events used in the computation.

Systematic uncertainties are computed to account for the different event selections used in the analysis. In particular, a $p_T$-binned FF is computed in a region with the
Figure 7. Distribution of $m_{\text{vis}}^{HH}$ in two regions enriched in multi-jet events, (a) ‘MJ OS 0-tag’ and (b) ‘MJ OS 1-tag’ in table 2. The background referred to as ‘Others’ contains $W$+jets, diboson, $ZH$, $t\bar{t}$ and single-top-quark events. The lower panels show a bin-by-bin comparison of the data and the predicted total background, in terms of event yield ratios. The hatched bands represent combined statistical and systematic uncertainties.

same selections as those used in ‘FF SS’, except that the charge requirement for the di-\(\tau\) candidate is changed to $Q = -1$. The deviation (up to 30\%) from the nominal FF is taken as a systematic uncertainty in the extrapolation from the SS region to the OS region. In addition, the maximum overall difference (42\%) between the FF values evaluated in regions with zero, one or two \(b\)-tagged track-jets in the selected large-\(R\) jet accounts for the systematic uncertainty in the event categorisation based on \(b\)-tagging.

The estimation of the multi-jet background is validated in two regions (‘MJ OS 0-tag’ and ‘MJ OS 1-tag’ in table 2) selected by requiring $Q = -1$ for the di-\(\tau\) candidate, either zero or one \(b\)-tagged track-jet in the selected large-\(R\) jet, and $|\Delta \phi| > 1$ between the selected di-\(\tau\) object and the missing transverse momentum to suppress events with correctly identified di-\(\tau\) objects. As shown in figure 7, good agreement between the predicted and measured multi-jet backgrounds is found in the ‘MJ OS 0-tag’ region, while some discrepancy is observed between the predicted and measured event yields in the ‘MJ OS 1-tag’ region.\(^3\) Therefore, an additional, conservative, non-closure systematic uncertainty of 50\% is assigned for the multi-jet background estimation in all regions with $N_{b\text{-tags}} \geq 1$.

\(^3\)The simulated background events with a true di-\(\tau\) object are corrected to account for the measured efficiency of the di-\(\tau\) tagger, as discussed in section 8, and an additional normalisation factor is subsequently applied to correct for the modelling of the $Z$+hf background, as discussed in section 9.
8 Data-driven correction of the di-τ tagger efficiency

To account for differences in the efficiency of the di-τ tagger between simulation and data, a correction is derived in a region enriched with properly identified di-τ objects from boosted $Z \to \tau^+\tau^-$ decays. This region, labelled as ‘Zττ 0-tag’ in table 2, is designed by imposing a veto on events with $b$-tagged track-jets in the selected large-$R$ jet and by demanding that $Q = -1$ and that the boosted di-τ system and the missing transverse momentum point in the same direction, i.e. $|\Delta \phi| < 1$. Figure 8 shows the visible di-τ mass in this region, as well as $p_T$ and $\eta$ distributions after an additional requirement on the visible di-τ mass, $30\text{ GeV} < m_{\text{vis}}^{\text{di-τ}} < 90\text{ GeV}$, which increases the fraction of $Z \to \tau^+\tau^-$ events to 87%.

A scale factor (SF) is computed in the ‘Zττ 0-tag’ region as the ratio of measured to predicted event yields with a true boosted hadronically decaying $\tau^+\tau^-$ pair. The measured event yield is obtained after subtracting the backgrounds with a misidentified di-τ object from the data:

$$\text{SF} = \frac{N(\text{data}) - N(\text{non-di-τ})}{N(\text{true di-τ})} = 0.84 \pm 0.09 \text{ (stat)} +^{0.14}_{-0.13} \text{ (Z-modelling)} +^{0.19}_{-0.20} \text{ (syst)}.$$  

No significant dependence of the ratio on the $p_T$ or $\eta$ of the di-τ object is found within the uncertainties, as illustrated in figures 8(b) and 8(c). Hence, a single-bin SF is considered in order to minimise the statistical uncertainty. The SF is applied to all simulated (background and signal) events containing a di-τ object matched to a hadronically decaying $\tau^+\tau^-$ pair at generator level.

The following uncertainties are considered in the SF computation. The statistical uncertainty accounts for the limited size of the data and simulated background samples, and is respectively 10.3% and 2.7% relative to the SF. The uncertainty labelled as ‘Z-modelling’ arises from the cross-section (5%) and the acceptance of $Z +$jets events. The latter comes from the choice of PDF set, as well as from variations of $\alpha_S$ and of the renormalisation, factorisation, matrix-element matching and soft-gluon resummation scales in Sherpa. The uncertainty related to the cross-section and acceptance of the other simulated processes is found to be negligible. The systematic uncertainties arising from the multi-jet background estimate are discussed in section 7. In all samples of simulated events, the instrumental systematic uncertainties arise primarily from the mismodelling of large-$R$ jets, in particular the jet energy scale and resolution, as well as the jet mass scale and resolution [99]. Uncertainties arising from the mismodelling of the reconstructed energy (momentum) measurements for electrons (muons), as well as from the mismodelling of their reconstruction and identification efficiencies [77, 78], are found to be negligible. The systematic uncertainties arising from the $E_{\text{T}}^{\text{miss}}$ scale and resolution [79], as well as from the mismodelling of pile-up, are taken into account. Finally, the uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [100], obtained using the LUCID-2 detector [101] for the primary luminosity measurements.

---

4The contribution of $Z \to \tau^+\tau^-$ events in which one of the τ-leptons decays to an electron or muon is negligible. A muon rarely deposits enough energy in the calorimeter to be reconstructed as a sub-jet of the di-τ object. On the other hand, the di-τ reconstruction and identification algorithms provide only little discrimination against $Z \to \tau^+\tau^-$ events with an electron in the decay chain. However, the lepton veto, which excludes non-isolated electrons, ensures that such events only have a minor contribution.
Figure 8. Distributions of (a) the visible mass of the di-τ object, (b) its $p_T$ and (c) $\eta$ in the region labelled as ‘$Z\tau\tau$ 0-tag’ in table 2. All simulated events containing a generator-level $\tau^+\tau^-$ pair matched to a simulated di-τ object are referred to as ‘True di-τ’. The lower panels show a bin-by-bin comparison of the data and the predicted total background, in terms of event yield ratios. The hatched bands represent combined statistical and systematic uncertainties.

The modelling of di-τ objects from genuine hadronically decaying $\tau^+\tau^-$ pairs is checked in the region ‘$Z\tau\tau$ 1-tag’ with exactly one $b$-tagged track-jet in the selected large-$R$ jet. In the following, for events containing a large-$R$ jet with $b$-tagged track-jets, an additional systematic uncertainty arises from differences between the predicted and measured $b$-tagging efficiency and rates at which $c$-jets and light-flavour jets are misidentified as $b$-tagged jets [93, 102, 103]. As shown in figure 9, good agreement between predicted and measured yields is observed in the region ‘$Z\tau\tau$ 1-tag’, within the uncertainties.
9 Search for resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ final state

The signal region (SR) used in the search for resonant Higgs boson pair production is defined by requiring OS charges for the two leading sub-jets of the di-$\tau$ object ($Q = -1$), $|\Delta\phi| < 1$, and two $b$-tagged track-jets in the selected large-$R$ jet, which furthermore must have a mass $m_J$ compatible with the mass of the Higgs boson (60–160 GeV). Finally, a requirement on $m_{\text{vis}}^{HH}$ is imposed, which depends on the resonance mass hypothesis: no requirement is applied if $m_X < 1.6$ TeV, whereas the cut-off value is set to 900 (1200) GeV if $m_X \geq 1.6$ (2.5) TeV. These requirements are summarised in the last row of table 2. The signal acceptance times efficiency at various stages of the event selection is shown in figure 10. The efficiency at low resonance masses is limited by the trigger thresholds and kinematic requirements on the large-$R$ jet.

In the SR, the dominant backgrounds arise from multi-jet, $ZH$, $Z+hf$ and $Z+lf$ events. The multi-jet background with a misidentified di-$\tau$ object is estimated with the data-driven method of section 7. The distribution shapes of events from the other background processes, with a true di-$\tau$ object, are modelled using simulation, but the data-driven SF computed in section 8 is employed for their normalisation. In addition, the SHERPA event generator was found to underestimate the heavy-flavour contribution in the $Z+\text{jets}$ process in previous ATLAS analyses, such as in ref. [104]. Therefore, the normalisation of the $Z \rightarrow \tau^+\tau^- + hf$ background is further corrected by a normalisation factor derived in a dedicated region enriched with $Z \rightarrow \ell^+\ell^- + hf$ events, where $\ell$ denotes either an electron or a muon defined as a ‘loose’ lepton in ref. [104]. This region is selected by requiring two isolated electrons or muons with an invariant mass $81 \text{ GeV} < m_{\ell\ell} < 101 \text{ GeV}$ and a dilepton transverse

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure9.png}
\caption{Predicted and measured event yields in the region labelled as ‘$Z\tau\tau$ 1-tag’ in table 2. All simulated events containing a generator-level $\tau^+\tau^-$ pair matched to a simulated di-$\tau$ object are referred to as ‘True di-$\tau$’. The main contribution to the background referred to as ‘Others’ comes from the $t\bar{t}$ process (about 3.6 events). The lower panel shows the event yield ratio of the data and the predicted total background. The hatched bands represent combined statistical and systematic uncertainties.}
\end{figure}
momentum $p_T^{\ell\ell} > 300$ GeV. In addition, the selected events must contain at least one large-$R$ jet with $p_T^J > 250$ GeV, $m_J > 50$ GeV and two associated $b$-tagged track-jets. After applying these event selections, the fraction of $Z$+hf events is about 80%. Figure 11 shows a comparison of the predicted and measured $p_T^{\ell\ell}$ distributions. After subtracting all other processes from the data and comparing with simulated $Z$+hf events, the normalisation factor is found to be $1.20 \pm 0.10 \text{(stat)} \pm 0.28 \text{(syst)}$. The statistical uncertainty accounts for the limited size of the data and simulated event samples, and is respectively 8.3% and 1.7% relative to the normalisation factor. Systematic uncertainties arise from the instrumental sources detailed in section 8 (3.4%) and from neglecting the differences between the $p_T$ distributions of the visible decay products of $Z \rightarrow \ell^+\ell^-$ and $Z \rightarrow \tau^+\tau^-$, due to the presence of neutrinos in the latter case (23%).

Table 3 summarises the event yields for data and all backgrounds, together with the associated statistical and systematic uncertainties. The systematic uncertainties in the multi-jet background estimation and of instrumental origin are the same as those discussed in sections 7 and 8. The uncertainty related to the cross-sections and acceptances of the simulated background processes is found to be negligible compared to the statistical uncertainty. None of the simulated $t\bar{t}$ events pass the event selection of the $HH$ SR, and hence a $+1\sigma$ uncertainty of 0.12 events in the associated yield is estimated by considering the simulated events that survive the selection criteria prior to applying requirements on $m_J$ and $m_{HH}^{vis}$.

The systematic uncertainties in the modelling of the signal acceptance and efficiency arise from four sources. Those due to missing higher-order corrections are estimated by changing the renormalisation and factorisation scales up and down by a factor of two. Systematic uncertainties arising from the strong coupling parameter are assessed by varying its value. Uncertainties related to the PDF set used for matrix-element generation are computed by comparing the signal acceptance and efficiency across the 100 replicas of the
Figure 11. Distribution of $p_T^{\ell\ell}$ in $Z \rightarrow \ell^+\ell^-$ events produced in association with heavy-flavour jets. The background labelled as ‘Others’ contains $W$+jets, diboson, $t\bar{t}$ and single-top-quark events. The lower panel shows a bin-by-bin comparison of the data and the predicted total background, in terms of event yield ratios. The hatched bands represent combined statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>Selection on $m^{vis}_{HH}$</th>
<th>&gt; 0 GeV</th>
<th>&gt; 900 GeV</th>
<th>&gt; 1200 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z\tau\tau$+hf</td>
<td>0.89 ± 0.25$^{+0.47}_{-0.35}$</td>
<td>0.75 ± 0.21$^{+0.47}_{-0.37}$</td>
<td>0.17 ± 0.05 ± 0.07</td>
</tr>
<tr>
<td>$Z\tau\tau$+lf</td>
<td>0.05 ± 0.05 ± 0.03</td>
<td>0.05 ± 0.05 ± 0.03</td>
<td>—</td>
</tr>
<tr>
<td>Multi-jet</td>
<td>0.18 ± 0.03 ± 0.14</td>
<td>0.17 ± 0.03 ± 0.13</td>
<td>0.09 ± 0.02 ± 0.07</td>
</tr>
<tr>
<td>$ZH$</td>
<td>0.11 ± 0.01 ± 0.04</td>
<td>0.09 ± 0.01 ± 0.03</td>
<td>0.02 ± — ± 0.01</td>
</tr>
<tr>
<td>Others</td>
<td>0.13 ± 0.05$^{+0.15}_{-0.07}$</td>
<td>0.13 ± 0.05$^{+0.15}_{-0.07}$</td>
<td>0.05 ± 0.03$^{+0.12}_{-0.03}$</td>
</tr>
<tr>
<td>Sum of backgrounds</td>
<td>1.36 ± 0.26$^{+0.42}_{-0.38}$</td>
<td>1.19 ± 0.23$^{+0.51}_{-0.40}$</td>
<td>0.33 ± 0.07$^{+0.16}_{-0.10}$</td>
</tr>
<tr>
<td>Data</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. Event yields of the various estimated backgrounds and data, computed in the SR of the search for $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$. The background labelled as ‘Others’ contains $W$+jets, diboson, $t\bar{t}$ and single-top-quark events. Statistical and systematic uncertainties are quoted, in that order. The background yields and uncertainties are pre-fit and are found to be similar to those post-fit. A dash ‘—’ indicates a yield or uncertainty below 0.01.

nominal NNPDF2.3LO PDF set, as well as with alternative PDF sets, MMHT2014LO [22] and CT14 [105]. Finally, the impact of the parton shower and hadronisation modelling is assessed by comparing the signal acceptance and efficiency between HERWIG 7 and PYTHIA 8 predictions. The latter is the dominant systematic uncertainty in the signal modelling (2–4% for most of the mass hypotheses).

Figure 12 shows the $m^{vis}_{HH}$ distribution in the SR, prior to the final selection using that variable. No significant excess with respect to the background expectation is found and only two events are observed, with $m^{vis}_{HH}$ values of 1012 and 1114 GeV.
In order to extract exclusion limits on the $X \to HH$ signal cross-section, a single-bin counting experiment is performed for every mass hypothesis, based on a profile-likelihood fit [106]. All sources of systematic uncertainty are accounted for as nuisance parameters that are profiled in the fit. The 95% CL upper limits on the $X \to HH$ cross-section are calculated as a function of $m_X$ using the CL$_s$ method [107]. Because both the expected and observed event yields are small, the distribution of the test statistic is sampled by generating pseudo-experiments. For each mass hypothesis, 10,000 background-only and 10,000 signal+background pseudo-datasets are produced, and the exclusion limits are obtained by scanning the signal strength and interpolating to find the value at which CL$_s$ = 0.05. The corresponding upper limits on the production cross-section of a heavy, narrow, scalar resonance decaying into a pair of Higgs bosons are shown in figure 13.

Two different requirements are applied to $m_{HH}^{\text{vis}}$ for the resonance mass hypotheses of 1.6 TeV and 2.5 TeV, leading to discontinuities in the limits (at 1.6 TeV, the difference between imposing no requirement and $m_{HH}^{\text{vis}} > 900$ GeV is less than 1% though). The more stringent requirement on $m_{HH}^{\text{vis}}$ reduces the expected background for resonance mass hypotheses beyond 2.5 TeV, and no data events survive. This leads to an observed exclusion limit slightly below the expected limit. For resonance masses in the range 1.2–3.0 TeV, the observed (expected) 95% CL upper limits lie between 94 and 28 fb (74 and 32 fb). For $m_X = 1.0$ TeV, where the analysis is least optimal, mostly due to the $p_T$ requirements on the large-$R$ jet at both the trigger and analysis levels, the observed (expected) limit becomes 817 fb (624 fb). These 95% CL upper limits are driven mainly by the statistical uncertainties of the data sample. The relative impact of the systematic uncertainties is estimated by comparing the expected upper limits when taking or not into account the nuisance parameters in the statistical analysis. For resonance mass hypotheses in the range 1.2–2.0 TeV, the expected upper limits deteriorate by 20–30%, mostly due to the
uncertainties in the di-τ reconstruction. At lower masses, the uncertainties in the large-$R$ jet energy and mass scales further affect the expected upper limits (the largest impact from all nuisance parameters reaches 85% at 1 TeV). On the other hand, systematic uncertainties are found to have a negligible impact (less than 10%) on the expected upper limits for mass hypotheses beyond 2.5 TeV.

10 Conclusion

In this paper, analysis techniques using large-radius jets and exploiting jet substructure are used in the ATLAS experiment to reconstruct and identify boosted di-τ systems. These can arise from the decay $H \rightarrow \tau^+\tau^-$ of a Higgs boson with a large Lorentz boost. A dedicated benchmark of the di-τ tagger is designed to search for resonant Higgs boson pair production in the high-mass regime, $X \rightarrow HH \rightarrow b\bar{b}\tau^+\tau^-$, where only final states with two hadronically decaying τ-leptons are considered. This benchmark corresponds to a 60% di-τ identification efficiency and a rejection factor of $10^4$ for the background of quark- or gluon-initiated jets misidentified as boosted di-τ objects. The efficiency of the di-τ tagger is determined using 139 fb$^{-1}$ of LHC proton-proton collision data at $\sqrt{s} = 13$ TeV.

The search for resonant Higgs boson pair production via gluon-gluon fusion is carried out using the same dataset. No significant deviations from the Standard Model predictions are observed, and 95% confidence-level upper limits are set on the production cross-section of a heavy, narrow, scalar resonance decaying into a pair of Higgs bosons. Assuming SM branching fractions for the Higgs boson, the observed (expected) limits on the $X \rightarrow HH$ production cross-section lie between 94 and 28 fb (74 and 32 fb) for resonance mass hypotheses in the range 1.2–3.0 TeV.
Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russia Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZS, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; ERC, ERDF, Horizon 2020, Marie Skłodowska-Curie Actions and COST, European Union; Investissements d'Avenir Labex, Investissements d'Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [108].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


[41] G. Luisoni, P. Nason, C. Oleari and F. Tramontano, $HW^\pm/HZ + 0$ and $1$ jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO, JHEP 10 (2013) 083 [arXiv:1306.2542] [INSPIRE].


[98] ATLAS collaboration, Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, JHEP 01 (2019) 030 [arXiv:1804.06174] [inSPIRE].


Also at Graduate School of Science, Osaka University, Osaka; Japan.

Also at Hellenic Open University, Patras; Greece.

Also at IJCLab, Université Paris-Saclay, CNRS/IN2P3, 91405, Orsay; France.

Also at Institucio Catalana de Recerca i Estudis Avancats, ICREA, Barcelona; Spain.

Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg; Germany.

Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen; Netherlands.

Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia; Bulgaria.

Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest; Hungary.

Also at Institute of Particle Physics (IPP); Canada.

Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku; Azerbaijan.

Also at Instituto de Fisica Teorica, IFT-UAM/CSIC, Madrid; Spain.

Also at Joint Institute for Nuclear Research, Dubna; Russia.

Also at Louisiana Tech University, Ruston LA; United States of America.

Also at Moscow Institute of Physics and Technology State University, Dolgoprudny; Russia.

Also at National Research Nuclear University MEPhI, Moscow; Russia.

Also at Physics Department, An-Najah National University, Nablus; Palestine.

Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg; Germany.

Also at The City College of New York, New York NY; United States of America.

Also at TRIUMF, Vancouver BC; Canada.

Also at Università di Napoli Parthenope, Napoli; Italy.

Also at University of Chinese Academy of Sciences (UCAS), Beijing; China.

* Deceased