ATLAS NOTE

ATLAS-CONF-2012-142

October 17, 2012

Performance of the Reconstruction and Identification of Hadronic τ Decays in ATLAS with 2011 Data

The ATLAS Collaboration

Abstract

This note describes updates to the reconstruction and identification algorithms developed to efficiently select hadronically decaying τ leptons ($\tau_{\text{had-vis}}$) in ATLAS. These updates are with respect to the baseline algorithms applied to the 2010 data. Two independent multivariate methods to discriminate $\tau_{\text{had-vis}}$ from QCD jets have been re-optimised. Another identification method to reject electrons mis-identified as $\tau_{\text{had-vis}}$ has also been re-optimised. For a signal identification efficiency of approximately 50%, rejection factors in the range 20–130 against QCD jets and approximately 300 against electrons are achieved. A new procedure has also been developed to reject mis-identified muons. For a signal identification efficiency of 96%, a muon rejection factor of approximately 2 is achieved. Four different "tag-and-probe" measurements of the $\tau_{\text{had-vis}}$ identification efficiency and the electron to $\tau_{\text{had-vis}}$ mis-identification probability based on 2011 data are presented. The measured $\tau_{\text{had-vis}}$ identification efficiencies are in agreement with those estimated using Monte Carlo simulation, within uncertainties of 4-5% for $\tau_{\text{had-vis}}$ $p_{\text{T}} > 22$ GeV.

1 Introduction

The τ leptons play an important role in the ATLAS physics program as they provide a useful signature in searches for the Standard Model Higgs boson and new phenomena in a wide range of theoretical models [1]. This note describes updates to the reconstruction and identification algorithms [2] developed to efficiently select hadronically decaying τ leptons (τ_{had}).

The τ lepton is very similar to an electron or a muon, but with a mass of 1776.82 ± 0.16 MeV [3], it is the only lepton heavy enough to decay into hadrons as well as into leptons. Due to its short lifetime, 2.9×10^{-13} seconds [3], the proper decay length of τ leptons is 87.11 μm. Thus they decay inside the LHC beam pipe and their decay products are used for identification.

The leptonic τ decay modes have the following branching fractions [3]: $\text{Br}(\tau \to e\nu_e\nu_\tau) = 17.8\%$ and $\text{Br}(\tau \to \mu\nu_\mu\nu_\tau) = 17.4\%$. In the remaining cases (64.8%) τ leptons decay hadronically, predominately into one or three charged pions, a neutrino and often additional neutral pions. There is also a contribution from Cabibbo-suppressed modes to kaons with a branching fraction of 2.9%. Hadronically decaying τ leptons are categorised by the number of charged decay products, observed as the number of tracks or "prongs". Hadronic 1-prong decays ($\tau_{\text{1-prong}}$) are the most common, 49.5% relative to all τ decays, followed by 3-prong decays ($\tau_{\text{3-prong}}$), 15.2%.

The reconstruction and identification of τ leptons at hadron colliders are challenging from the experimental point of view. Since purely leptonic τ decays are very difficult to distinguish from prompt electrons or muons, the τ identification algorithms are developed to efficiently reconstruct and identify the visible part (without the neutrino) of the hadronic decay modes, referred by the term $\tau_{\text{had-vis}}$ in this document. The challenge when identifying hadronic τ decays is that their signatures in the detector are very similar to quark- or gluon-initiated jets from QCD processes (referred to as QCD jets). In addition, these background processes have cross sections many orders of magnitude greater than the cross sections for weak interaction processes involving τ leptons. The most discriminating features for identifying $\tau_{\text{had-vis}}$ from this multi-jet background are the τ lepton's characteristic 1- or 3-prong signature and the relatively narrow clustering of tracks and energy deposits in the calorimeters. Electrons and muons can also be misidentified as $\tau_{\text{1-prong}}$. Separate procedures have been developed for rejecting electrons and muons as their signatures are different from QCD jets.

After the initial observation of hadronically decaying τ leptons in ATLAS [4, 5], there has been rapid development of analyses featuring τ leptons, including the $W \to \tau\nu$ and $Z \to \tau\tau$ cross section measurements [6, 7], the τ polarisation measurement [8] as well as the searches for the Higgs boson [9–11] and supersymmetry [12, 13]. These studies have increased demand for improved τ reconstruction and identification performance. The reconstruction and identification of τ leptons have been re-optimised compared with that described in Refs. [2, 14, 15]. The main focus of the re-optimisation was to make the τ reconstruction less sensitive to changes in the number of overlapping proton-proton interactions per bunch crossing (pile up), which increased significantly in the 2011 data runs and reached the maximum average of 17 interactions per beam crossing. Two independent identification methods that discriminate τ leptons from QCD jets are discussed: a log-likelihood method (LLH) and a method using boosted decision trees (BDT). Another identification method has also been re-optimised to reject electrons mis-identified as $\tau_{\text{had-vis}}$ leptons using BDT as discrimination. In addition, a new procedure has been developed to reject muons mis-identified as $\tau_{\text{had-vis}}$ candidates.

This note also describes four different "tag-and-probe" measurements of the $\tau_{\text{had-vis}}$ identification efficiency and of the electron to $\tau_{\text{had-vis}}$ mis-identification probability on 2011 data. The electron to $\tau_{\text{had-vis}}$ mis-identification probability is measured using $Z \to ee$ events. The $\tau_{\text{had-vis}}$ identification efficiency measurements are based on $Z \to \tau\tau$ and $W \to \tau\nu$ events and use different background estimations. The measured efficiencies are compared to the ones estimated by Monte Carlo simulations and data/MC correction factors are then extracted. The Monte Carlo simulations used for the W and Z events in the

"tag-and-probe" measurements are produced with the ALPGEN [16] generator. These events are compared to simulations made with the PYTHIA [17] generator for the systematic uncertainty estimations. The default physics list for the signal simulation is `QGSP_BERT` [18]. All simulated events are passed through the full simulation of the ATLAS detector using GEANT4 [19] and are reconstructed with the same software as used for data.

2 Reconstruction of Hadronic Tau Decays

The $\tau_{\text{had-vis}}$ reconstruction algorithm relies on the inner detector and calorimeter information [14]. The inner detector provides information on the single or collimated track system. The charge of the decaying τ lepton can be directly determined from the charge(s) of its associated track(s). Particular attention has been given to minimising the amount of charge mis-identification and of migration between the n-prong categories during the reconstruction (Section 2.2). Calorimetry provides information on the energy deposits from the visible decay products. Hadronically decaying τ leptons are well-collimated resulting in a relatively narrow shower in the electromagnetic calorimeter with, for 1-prong decays with one or few π^0's, a significant electromagnetic component. The calorimeter and tracking information should match, with narrow calorimeter clusters being found near the track(s) impact point in the calorimeter.

The $\tau_{\text{had-vis}}$ reconstruction and identification algorithms in ATLAS follow these steps. The $\tau_{\text{had-vis}}$ reconstruction is "seeded" from reconstructed jets by considering each jet as a $\tau_{\text{had-vis}}$ candidate. The list of calorimeter clusters associated with each $\tau_{\text{had-vis}}$ candidate is then refined and used to calculate kinematic quantities. Tracks satisfying dedicated selection criteria are associated with the calorimeter clusters. A list of identification variables is then calculated from the tracking and calorimetry information. These variables are combined into multivariate discriminants to reject fake candidates from QCD jets and electrons. Finally, selection on the output of the discriminants is used at the analysis level to select a sample of $\tau_{\text{had-vis}}$ candidates with the desired level of background rejection and signal efficiency.

2.1 The $\tau_{\text{had-vis}}$ Reconstruction

The $\tau_{\text{had-vis}}$ reconstruction algorithm is seeded from jets reconstructed using the anti-k_t algorithm [20], with a distance parameter $R = 0.4$. Topological clusters [21] made of calorimeter cells calibrated with the Local Hadron Calibration (LC) [22] are used as an input for the jet algorithm. All jets with transverse momentum $p_{\text{T}} > 10$ GeV and pseudorapidity $|\eta| < 2.5$, which corresponds to the η-coverage of the ATLAS tracking system, seed the $\tau_{\text{had-vis}}$ reconstruction algorithm.

The reconstructed four-momentum of the $\tau_{\text{had-vis}}$ candidate is defined in terms of three degrees of freedom[1]: p_{T}, η, and ϕ. The η and ϕ are taken from the seed jet, which are determined by calculating the sum of the four-vectors of the constituent topological clusters, assuming zero mass for each of the constituents [23]. The mass of the $\tau_{\text{had-vis}}$ candidate is defined to be zero and consequently the transverse momentum, p_{T}, and the transverse energy, $E_{\text{T}} = E \sin\theta$, are identical. Because hadronic τ decays consist of a specific mixture of charged and neutral pions, the energy scale of hadronic τ candidates is calibrated independent of the jet energy scale. The reconstructed energy of $\tau_{\text{had-vis}}$ candidates is corrected to the final energy scale by a Monte Carlo based calibration procedure using clusters, within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.2$ of the seed jet barycentre axis. The $\tau_{\text{had-vis}}$ energy calibration is described in detail in a separate document in Ref. [24]. The final η position of the $\tau_{\text{had-vis}}$ candidate is also determined by

[1]The ATLAS Coordinate System [14] is a right-handed system with the x-axis pointing to the centre of the LHC ring, the y-axis pointing upwards, and the z-axis following the beam line. The spherical coordinates ϕ and θ are defined in the usual way, with the azimuthal angle, ϕ, measuring the angle in the xy-plane from the positive x-axis, increasing towards positive y. θ measures the angle from the positive z-axis, but this coordinate is often specified by the pseudorapidity, η, defined as $\eta = -\ln(\tan\frac{\theta}{2})$.

the energy calibration procedure, which takes into account clusters reconstructed in poorly instrumented regions of the calorimeter [24].

2.2 Track Association

Tracks are associated with each $\tau_{\text{had-vis}}$ candidate if they are within the *core cone*, defined as the region within $\Delta R < 0.2$ of the axis of the seed jet, and satisfy the following quality criteria based on the inner detector pixel and SCT silicon detectors:

- $p_{\text{T}} > 1$ GeV,
- Number of pixel hits ≥ 2,
- Number of pixel hits + number of SCT hits ≥ 7,
- $|d_0| < 1.0$ mm,
- $|z_0 \sin\theta| < 1.5$ mm,

where d_0 is the distance of closest approach of the track to the reconstructed primary vertex in the transverse plane, while z_0 is the longitudinal distance of closest approach.

As the number of interaction vertices per beam crossing, or pile up, increases, the ability for general tracking to correctly identify the primary vertex degrades when using the "default" primary vertex, which is defined as the primary vertex candidate with the highest $\sum(p_T^{\text{trk}})^2$ [25]. The $\tau_{\text{had-vis}}$ track association, in particular the $|z_0 \sin\theta|$ requirement, is very sensitive to the selected primary vertex, which in turn affects both $\tau_{\text{had-vis}}$ reconstruction and identification. A new algorithm used to mitigate these effects is presented in Section 2.3.

The $\tau_{\text{had-vis}}$ candidates are classified as *n*-prong depending on the number *n* of tracks counted in the core cone. Multi-prong refers to $\tau_{\text{had-vis}}$ candidates with more than one track. Tracks within the *isolation annulus*, defined by $0.2 < \Delta R < 0.4$ of the axis of the seed jet, are also counted for variable calculations and are required to satisfy the same track quality criteria.

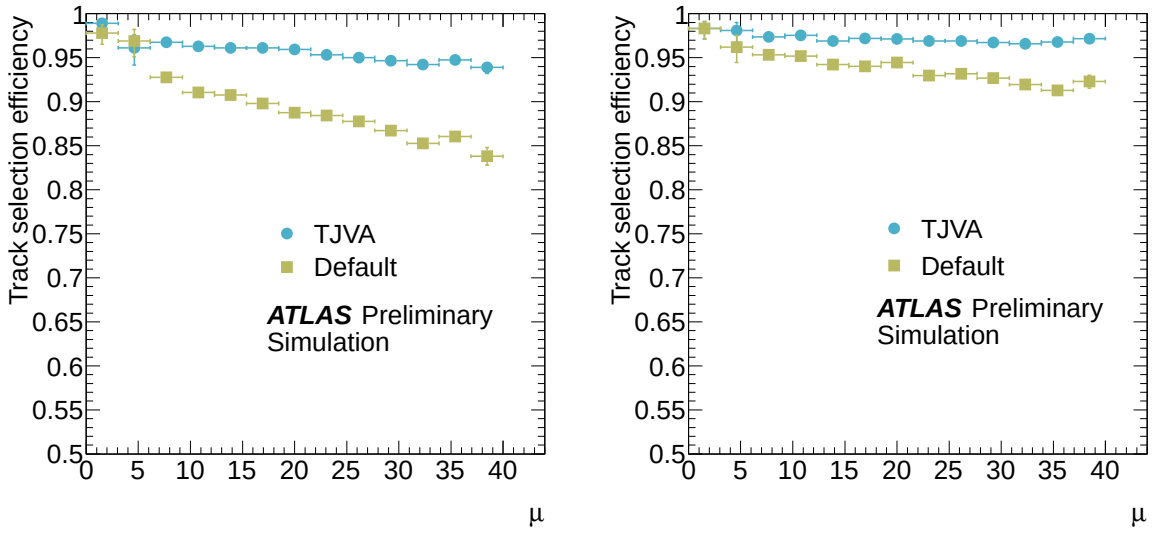
2.3 Vertex Association

The probability of incorrectly assigning a pile up vertex as the primary vertex has increased with the larger number of pile up collisions in 2011. This causes tracks to fail the z_0 impact parameter requirement, as observed in simulated $Z \rightarrow \tau\tau$ events. Figures 1(a) and 1(b) demonstrate the $\tau_{\text{had-vis}}$ track selection efficiency at different pile up conditions, where $\tau_{\text{had-vis}}$ track selection efficiency is defined as the probability for a real charged pion from a τ_{had} decay to be associated with a reconstructed $\tau_{\text{had-vis}}$ candidate.

As illustrated in Figures 1(a) and 1(b), the $\tau_{\text{had-vis}}$ track selection performance is degraded with increased pile up conditions. These inefficiencies affect the number of tracks associated with the reconstructed $\tau_{\text{had-vis}}$ objects as well as the calculation of many variables used by the $\tau_{\text{had-vis}}$ identification algorithms (Sections 2.4 and 3.1).

To ensure that the $\tau_{\text{had-vis}}$ track association method is robust with respect to pile up conditions, a new algorithm has been developed to correctly identify the primary vertex hypothesis for each $\tau_{\text{had-vis}}$ candidate. This new algorithm, called Tau Jet Vertex Association (TJVA), is built using the existing Jet Vertex Association (JVA) algorithm. The JVA algorithm works by finding for each jet a vertex candidate with the highest Jet Vertex Fraction (f_{JVF}), calculated using the following formula:

$$f_{\text{JVF}}(\text{jet}|\text{vtx}) = \frac{\sum p_T^{\text{trk}|\text{vtx}}}{\sum p_T^{\text{trk}}}$$



(a) True $\tau_{1\text{-prong}}$.

(b) True $\tau_{3\text{-prong}}$.

Figure 1: The $\tau_{\text{had-vis}}$ track selection efficiency with respect to the average number of pile up interactions per bunch crossing (μ) for reconstructed $\tau_{\text{had-vis}}$ candidates in $Z \rightarrow \tau\tau$ simulated events. Only $\tau_{\text{had-vis}}$ candidates with $p_T > 15$ GeV matching to a truth-tau within $\Delta R < 0.2$ are considered. With TJVA, the $\tau_{\text{had-vis}}$ track multiplicity is less sensitive to pile up and a smaller degradation in efficiency is observed as pile up increases.

where f_{JVF}(jet|vtx) is the jet vertex fraction of a jet given a vertex candidate, "trk|vtx" refers to a track matched to a given vertex and "trk" in the denominator refers to tracks in the jet. This track-vertex match can be customised using the impact parameter and the longitudinal distance. Moreover, the set of tracks used in the calculation of f_{JVF} can be customised by applying selection criteria on the track kinematics and quality criteria. Figure 2 illustrates the JVA algorithm.

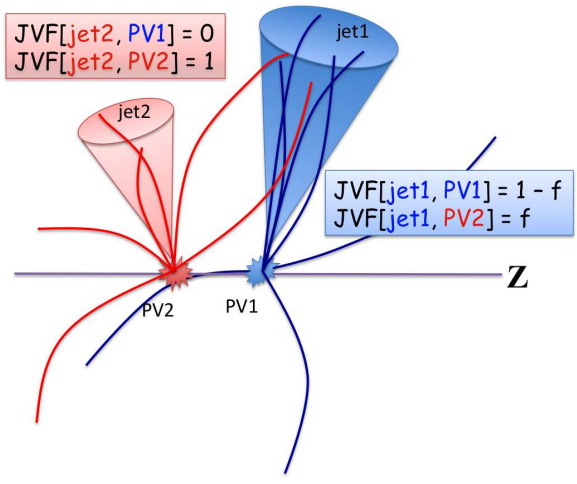
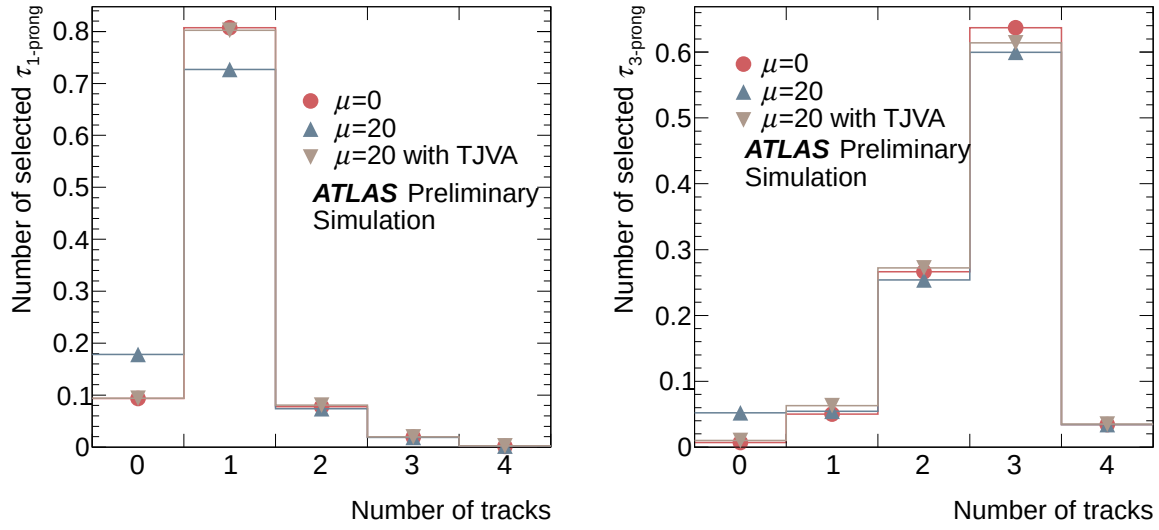


Figure 2: Depiction of the jet-vertex fraction discriminant.

TJVA uses Tau Jet Vertex Fraction (f_{TJVF}), which is calculated similarly to f_{JVF}. The track-vertex match criteria are the same as for f_{JVF} but the track selection parameters for f_{TJVF} have been optimised

for tracks from τ_{had} decays. These optimised track selection criteria are the same criteria described in Section 2.2, except for the impact parameter requirements that are not used for this track selection. It is important to stress that these selection criteria are only used to calculate f_{TJVF}, and thus, in the choosing of a primary vertex candidate for the $\tau_{\text{had-vis}}$ candidate. Once the new impact parameters are recalculated with respect to the TJVA primary vertex candidate, $\tau_{\text{had-vis}}$ tracks are selected using the criteria described in Section 2.2, which include the impact parameter requirements.

Figures 3(a) and 3(b) show the corrected $\tau_{\text{had-vis}}$ track multiplicity, where it can be seen that a large fraction of its degradation due to pile up interactions can be recovered by the TJVA algorithm, especially for $\tau_{\text{1-prong}}$ decays.



(a) True $\tau_{\text{1-prong}}$.

(b) True $\tau_{\text{3-prong}}$.

Figure 3: Distribution of the number of tracks associated with a reconstructed $\tau_{\text{had-vis}}$ candidate in $Z \to \tau\tau$ simulated events with no pile up, $\mu = 20$, and $\mu = 20$ with TJVA. Only $\tau_{\text{had-vis}}$ candidates with $p_T > 15$ GeV matching to a true τ_{had} within $\Delta R < 0.2$ are considered and the distributions are normalised to unity. With TJVA, the $\tau_{\text{had-vis}}$ track multiplicity is less sensitive to pile up, especially in $\tau_{\text{1-prong}}$ decays.

2.4 Reconstructed Variables

The reconstruction of $\tau_{\text{had-vis}}$ candidates provides very little rejection against the QCD jet background to hadronically decaying τ leptons. This rejection is provided in a separate identification step, using discriminating variables that are calculated during the reconstruction. These variables are defined in Appendix A and are used for identification as discussed in Section 3.

3 Identification of Hadronic Tau Decays and Lepton Veto

3.1 Algorithms for $\tau_{\text{had-vis}}$ Identification

Two independent $\tau_{\text{had-vis}}$ identification methods using boosted decision trees (BDT) and log-likelihood (LLH) as discriminants are discussed in the following. The set of variables used by each method remain unchanged with respect to previous definitions [15].

The $\tau_{\text{had-vis}}$ candidates from $Z \to \tau\tau$, $Z' \to \tau\tau$ and $W \to \tau\nu$ simulated events and background data samples enriched with QCD jets are used for the re-optimisation of these $\tau_{\text{had-vis}}$ identification algorithms. $Z' \to \tau\tau$ events with Z' masses of 500 GeV, 750 GeV, 1000 GeV and 1250 GeV are included to enhance the number of $\tau_{\text{had-vis}}$ candidates with higher p_{T}. The background events are data events recorded in 2011 by applying a QCD di-jet event selection: two high p_{T} back-to-back jets ($\Delta\phi(\text{jet}_1, \text{jet}_2) > 2.7$, $p_T^{\text{jet}_1} > 30$ GeV and $p_T^{\text{jet}_2} > 15$ GeV) in which the jet with the highest transverse momentum is used to tag the event and the other one is used as the mis-identified $\tau_{\text{had-vis}}$ candidate.

The $\tau_{\text{had-vis}}$ identification efficiency is defined as the fraction of the number of identified 1-prong (multi-prong) $\tau_{\text{had-vis}}$ candidates matching to a true $\tau_{\text{1-prong}}$ ($\tau_{\text{3-prong}}$) over the number of true $\tau_{\text{1-prong}}$ ($\tau_{\text{3-prong}}$); true τ_{had} with $p_T > 20$ GeV and $|\eta| < 2.5$ are considered and $\Delta R < 0.2$ is used as matching criteria. The background efficiency, measured with the QCD di-jet data sample, is defined as the number of probe jets reconstructed and identified as a $\tau_{\text{had-vis}}$ candidate over the number of probe jets reconstructed as $\tau_{\text{had-vis}}$ candidates.

3.2 Boosted Decision Trees Identification

The boosted decision tree identification method remains unchanged from the identification previously documented in Ref. [15]. To compensate for small changes in the distributions of the input variables and for higher pile up conditions, only the criteria on the BDT score corresponding to the *loose*, *medium*, and *tight* signal efficiency working points have been adjusted. For 1-prong candidates these working points are signal efficiencies of 60%, 50% and 30% and for multi-prong candidates they are 65%, 55% and 35%. The higher multi-prong efficiency compensates for the lower relative contribution of 3-prong $\tau_{\text{had-vis}}$ in the multi-prong category. Due to the dependence of the BDT score distributions on the p_{T} of the $\tau_{\text{had-vis}}$ candidates, these working points must be determined as functions of the reconstructed $\tau_{\text{had-vis}}$ p_{T} to yield approximately constant efficiency over p_{T}. Reconstructed $\tau_{\text{had-vis}}$ candidates from signal events are split into 20 variable-width p_{T} bins from 20 GeV to 100 GeV containing approximately an equal number of $\tau_{\text{had-vis}}$. Requirements on the BDT score which yield the target signal efficiencies are then determined separately in each bin for 1 and multi-prong candidates. The threshold values are then interpolated between bin centres and are constant above 100 GeV.

To provide working points that are also approximately independent of the pile up conditions, the above procedure is performed in separate bins of the number of vertices ($n_{\text{vtx}} \leq 3$; $3 < n_{\text{vtx}} \leq 5$; $5 < n_{\text{vtx}} \leq 7$ and $n_{\text{vtx}} > 7$).

3.3 Log-likelihood Identification

The log-likelihood method used for this optimisation is based on the previous one described in Ref. [15]. No re-training was performed, i.e. the same probability density functions were used. Similarly to Section 3.2, the threshold value on the likelihood score is calculated as a function of p_{T} to obtain a constant signal efficiency as a function of the transverse momentum of the $\tau_{\text{had-vis}}$ candidate. This procedure accounts for the p_{T} dependence of the input variables and hence of the score itself. Furthermore, the target efficiencies defining the *loose*, *medium* and *tight* working points have been adjusted to yield on average 60%, 50% and 30% signal efficiency for 1-prong $\tau_{\text{had-vis}}$ candidates and 65%, 55% and 35% for multi-prong $\tau_{\text{had-vis}}$ candidates, respectively. The threshold values are adjusted to also yield a flat signal efficiency as a function of the number of reconstructed vertices, n_{vtx}.

3.4 τ_{had} Identification Results

The signal efficiency as a function of the the number of reconstructed vertices is summarised in Figure 4 for both identification methods. As expected, the efficiency does not have a strong dependence on the

number of vertices.

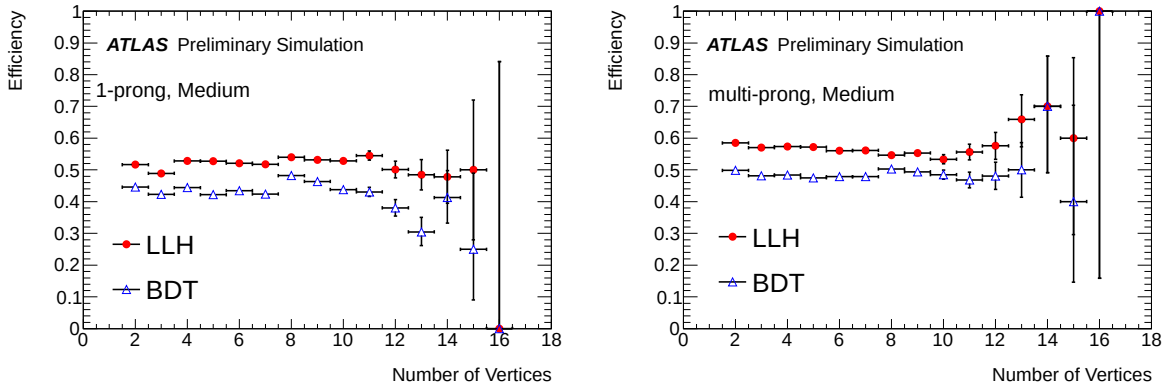


Figure 4: Signal efficiency for 1-prong (left) and multi-prong (right) for *medium* identification working point as a function of the number of reconstructed vertices in the event. Differences in the actual efficiencies shown here and the target efficiencies arise from differences in pile up conditions that were simulated between the samples used to determine the *loose*, *medium*, and *tight* cuts on the discriminant output, and the Monte Carlo samples used to evaluate the $\tau_{\text{had-vis}}$ identification performance throughout this note.

The performance of the $\tau_{\text{had-vis}}$ identification methods are illustrated in Figure 5, in which the inverse background efficiency as a function of the signal efficiency for 1-prong and multi-prong $\tau_{\text{had-vis}}$ candidates are shown.

3.5 Electron Veto

The characteristic signatures of hadronically decaying 1-prong τ leptons can be mimicked by electrons. Despite the similarity of τ lepton and electron signatures, there are properties that can be used to distinguish between them. For example, the electromagnetic shower produced by a τ lepton in the calorimeter, which tends to be longer and wider than an electron-induced shower. These properties can be used to define $\tau_{\text{had-vis}}$ identification discriminants specialised in rejecting electrons mis-identified as $\tau_{\text{had-vis}}$ candidates. In the following, the discriminant using boosted decision trees (electron BDT) is described. Three working points: *loose*, *medium* and *tight*, corresponding to efficiencies of 95%, 85% and 75%, respectively, are optimised for the electron BDT discriminants. The signal efficiency used for the performance evaluation is defined as the fraction of reconstructed 1-prong $\tau_{\text{had-vis}}$ candidates matching a true 1-prong $\tau_{\text{had-vis}}$ passing loose cut-based $\tau_{\text{had-vis}}$ identification[2] that also satisfies the electron BDT discriminant.

The electron BDT is optimised using simulated $Z \to \tau\tau$ events for signal and simulated $Z \to ee$ events for background. The signal candidates are required to match to a true $\tau_{\text{1-prong}}$ lepton and background candidates are required to match to a true electron, both within $\Delta R < 0.2$. All candidates are required to have $p_{\text{T}} > 20$ GeV.

The electron BDT discriminant is performed in four regions of $|\eta|$: barrel ($|\eta| < 1.37$), crack ($1.37 < |\eta| < 1.52$), endcap ($1.52 < |\eta| < 2.0$) and forward endcap ($2.0 < |\eta| < 2.3$). The best performing and best modelled variables in each $|\eta|$ region are used for training. The variables are listed in Appendix A. The following variables illustrated in Figure 6 were added since the previous version of the electron BDT training: the electromagnetic energy over track momentum (f_{P}^{EM}), the calorimeter presampler strip

[2] Another $\tau_{\text{had-vis}}$ identification algorithm described in Ref. [2] that has not been re-optimised.

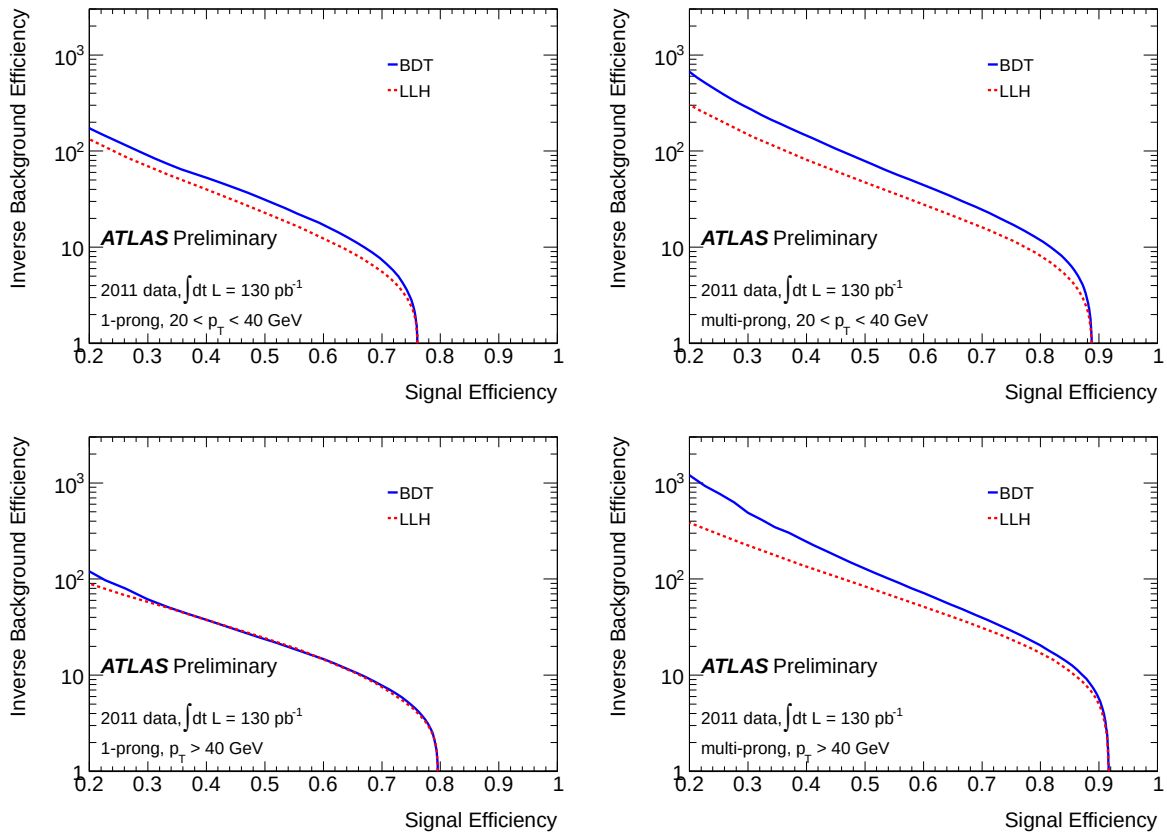


Figure 5: Inverse background efficiency versus signal efficiency for all discriminants on 1-prong (left) and multi-prong (right) $\tau_{\text{had-vis}}$ candidates. The top row shows the low p_{T} range, 20 GeV $< p_{\text{T}} \leq$ 40 GeV. The bottom row shows high p_{T}, i.e. $p_{\text{T}} >$ 40 GeV. The *loose*, *medium* and *tight* working points have 60%, 50% and 30% signal efficiency for 1-prong $\tau_{\text{had-vis}}$ candidates and 65%, 55% and 35% for multi-prong $\tau_{\text{had-vis}}$ candidates.

energy fraction (f_{PS}) and the ratio of electromagnetic energy of charged pions over calorimetric electromagnetic energy ($f_{\text{EM}}^{\pi^{\pm}}$). These variables are defined in functional form in Appendix A. The output score of the electron BDT discriminant for electrons and $\tau_{\text{1-prong}}$ selected from simulated events is shown in Figure 7. The electron BDT output score demonstrates good separation between electrons and $\tau_{\text{1-prong}}$. Three working points: *loose*, *medium*, and *tight* are defined to be a electron BDT score greater than 0.465, 0.565 and 0.645, respectively.

The efficiency for MC simulated signal τ candidates versus the electron rejection efficiency is shown in Figure 8. The efficiencies are shown for the electron BDT discriminant in central ($|\eta| < 2.0$) and forward endcap ($|\eta| > 2.0$) regions. The figures were produced with 1-prong reconstructed $\tau_{\text{had-vis}}$ candidates truth-matched to either electrons or $\tau_{\text{had-vis}}$ candidates, with $p_{\text{T}} >$ 20 GeV and identified by the *loose* cut-based $\tau_{\text{had-vis}}$ identification algorithm.

3.6 Muon Veto

As minimum ionising particles, muons do not deposit a significant amount of energy in the calorimeters and therefore a muon is unlikely to be reconstructed as a $\tau_{\text{had-vis}}$ candidate. However, when a cluster of energy in the calorimeter is associated with a muon, the muon track and the calorimeter cluster objects can be mis-identified as a $\tau_{\text{had-vis}}$. Therefore, a dedicated cut-based muon veto has been developed to

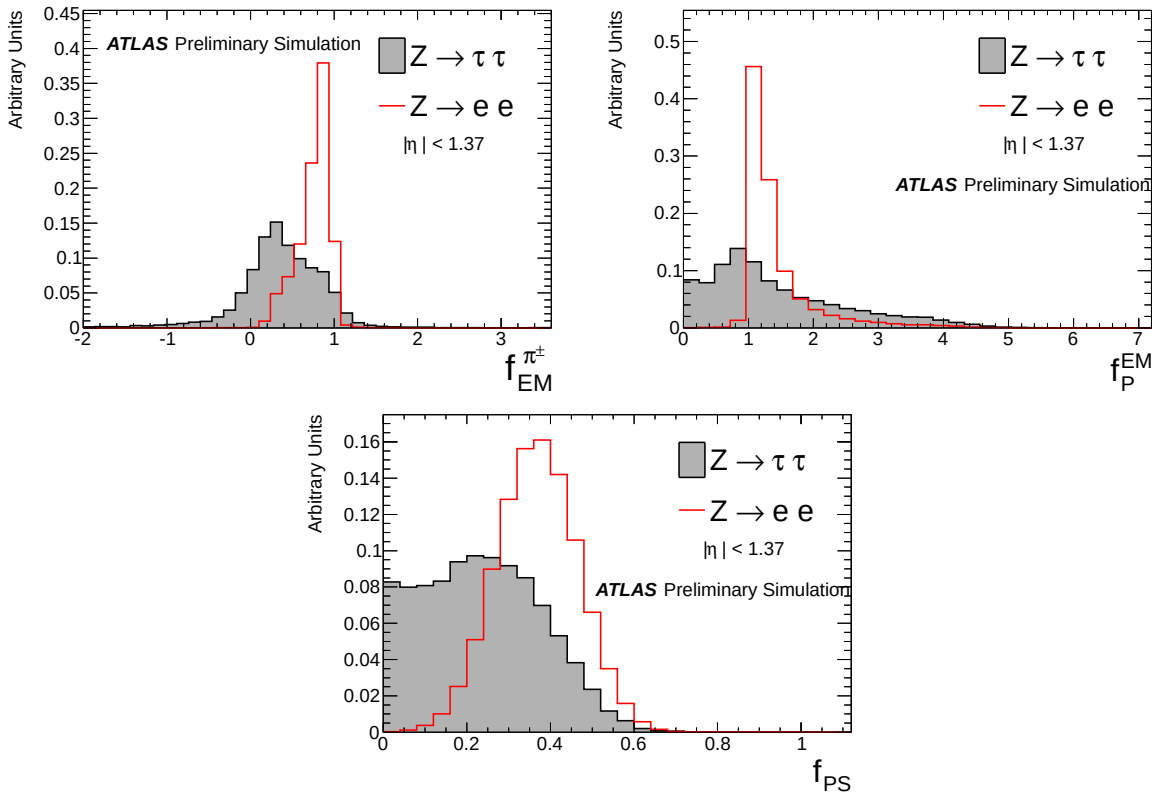


Figure 6: The distributions of the three new variables used in this current version of the electron BDT training, for the barrel region of ATLAS ($|\eta| < 1.37$).

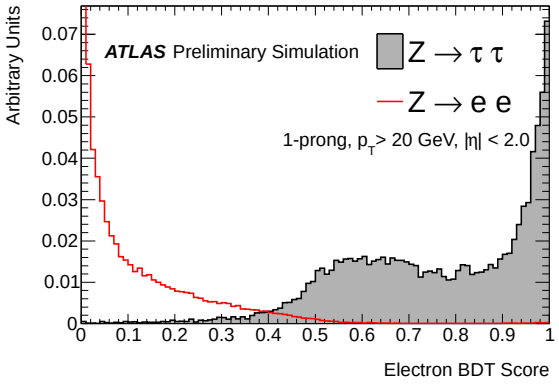


Figure 7: Score of the BDT-based electron veto for MC simulated $\tau_{\text{1-prong}}$ and electrons both reconstructed as 1-prong $\tau_{\text{had-vis}}$ candidates with $|\eta| < 2.0$.

reduce backgrounds from muon fakes, as described here.

Muons mis-reconstructed as $\tau_{\text{had-vis}}$ candidates can be classified according to the source of the associated calorimeter clusters:

- Case 1: The muon itself leaves anomalously large energy deposits in the calorimeter.
- Case 2: The energy comes from elsewhere, either collinear radiation from the muon prior to entering the calorimeter, or coincidental overlap with some other calorimeter clusters.

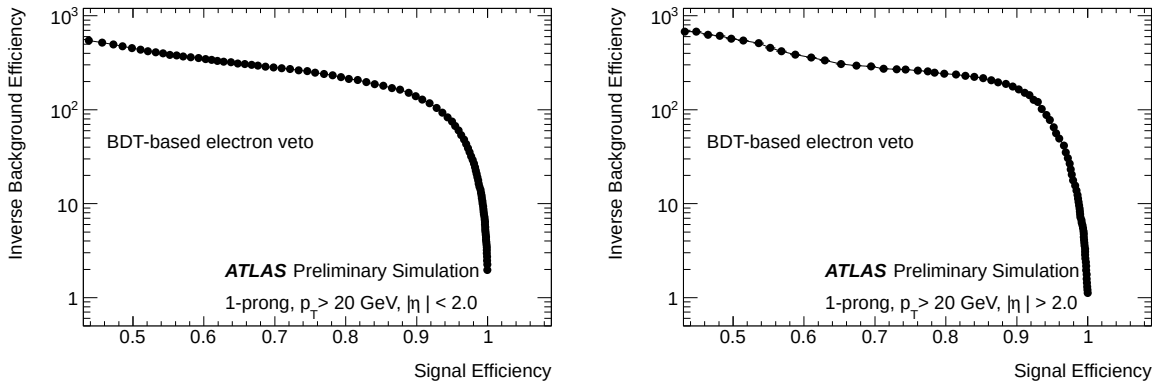


Figure 8: Inverse background (electron) efficiency as a function of signal efficiency for 1-prong reconstructed $\tau_{\text{had-vis}}$ candidates with $p_{\text{T}} > 20$ GeV, in the central ($|\eta| < 2.0$) and forward endcap ($|\eta| > 2.0$) regions, for the electron BDT discriminant.

The two cases can be clearly distinguished by the different shower shapes in the calorimeter: in Case 1, the muon will typically pass through the electromagnetic calorimeter, and so most of the energy will be deposited in the much deeper hadronic calorimeter; while in Case 2, the radiation is mostly electromagnetic and leakage into the hadronic calorimeter is minimal. Figure 9 shows the fraction of transverse energy of the τ_{had} candidate deposited in the electromagnetic calorimeter, f_{EM}, with the peaks at low and high values corresponding to Case 1 and Case 2, respectively. Additionally, in both cases there is no match between the track momentum and the calorimeter energy, so that the lead track momentum fraction (f_{track}) may be much higher than expected for true τ_{had}. Figure 10 shows this variable for the low and high f_{EM} regions.

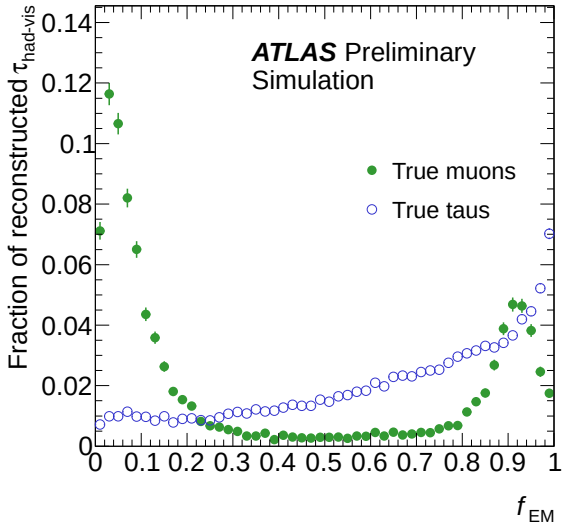
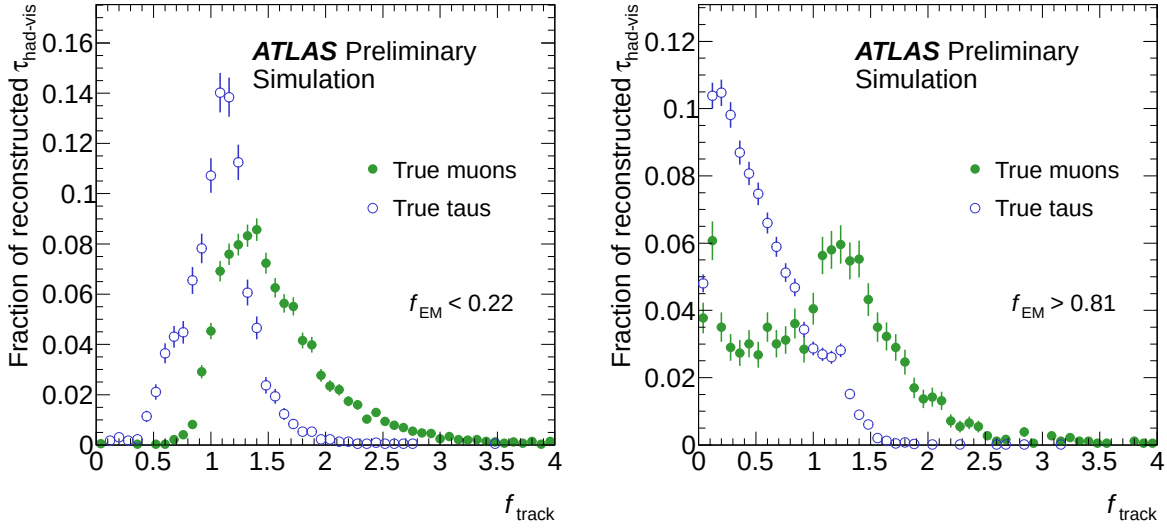


Figure 9: The electromagnetic fraction for reconstructed $\tau_{\text{had-vis}}$ candidates matched to true muons and to true τ_{had}.

To optimise the muon veto algorithm, true τ_{had} and muons with $p_T > 20$ GeV from simulated $Z \to \tau\tau$ and $Z \to \mu\mu$ samples were used. Because the muon veto is normally applied to muons that were not identified by the muon identification algorithm [1], an overlap removal was performed with respect



(a) $f_{EM} < 0.22$ (b) $f_{EM} > 0.81$

Figure 10: Lead track momentum fraction (f_{track}) for reconstructed $\tau_{had\text{-}vis}$ candidates matched to true muons and to true τ_{had}, in the low f_{EM} (left) and high f_{EM} (right) regions.

to muons identified by the muon identification algorithm with the *tight* working point. Objects are first classified into *low*, *medium*, and *high* f_{EM} regions, with medium f_{EM} objects automatically passing the veto. In the low and high f_{EM} regions, separate upper thresholds on f_{track} are then applied. These four parameters (the bounds on the low and high f_{EM} regions and the f_{track} thresholds in each region) were optimised to give maximum muon rejection for a target $\tau_{had\text{-}vis}$ efficiency of 96%. Both efficiency and rejection are defined with respect to BDT *medium* $\tau_{had\text{-}vis}$ identification, the *loose* cut-based electron veto and muon overlap removal. The efficiency is the fraction of such candidates passing the muon veto, and the rejection is the fraction of candidates failing the veto. The resulting parameter values are 0.22 and 0.81 for the boundaries of the f_{EM} regions, and 1.2 and 1.4 for the f_{track} threshold in the low and high regions, respectively. Approximately 55% of muons mis-identified as $\tau_{had\text{-}vis}$ are rejected by this veto; however, the performance can depend strongly on the details of the overlap removal performed by an analysis.

4 Measurement of the Electron to $\tau_{had\text{-}vis}$ Fake Rate

This section describes the measurement of the electron mis-identification probability with $Z \to ee$ events in 2011 data. The electron mis-identification probability is measured using different combinations of working points for jet and electron discrimination and different types of overlap removals between the $\tau_{had\text{-}vis}$ candidates and reconstructed electrons. The events used are a very pure sample of electrons that are reconstructed as $\tau_{had\text{-}vis}$ and they are selected with a tag-and-probe method in $Z \to ee$ events. The tagging object is an electron passing the tight electron identification [26] and the probe is a reconstructed $\tau_{had\text{-}vis}$ with the tag-probe invariant mass in the range [80, 100] GeV. The minimum transverse momentum required for the probe $\tau_{had\text{-}vis}$ is 20 GeV. Such a measurement has already been performed on early 2011 data and described in more detail in Ref. [27]. The background events are mainly QCD multi-jet events and they are estimated with a two-dimensional side-band extrapolation method, which assumes that the charge product between the tag electron and the probe $\tau_{had\text{-}vis}$ and the number of tracks of the probe $\tau_{had\text{-}vis}$ are not correlated. After the event selection, a good agreement is observed between data

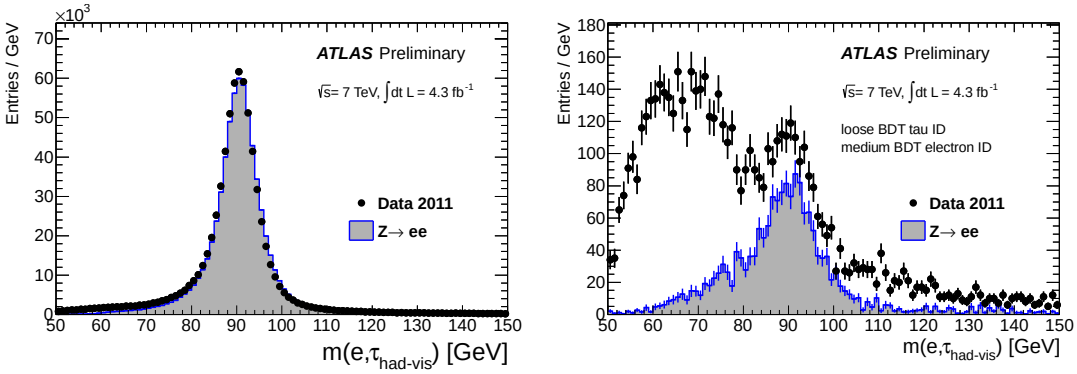


Figure 11: Invariant mass of the electron-$\tau_{\text{had-vis}}$ pair in the selected events. On the left plot, no discrimination is applied on the probe $\tau_{\text{had-vis}}$. The right plot shows only those events where the probe $\tau_{\text{had-vis}}$ has passed the BDT *loose* $\tau_{\text{had-vis}}$ identification and BDT *medium* electron veto.

and $Z \rightarrow ee$ Monte Carlo simulations even without any background subtraction, as shown in Figure 11 (left). However, after the application of the jet and the electron discrimination, the purity of electron events in data is significantly reduced, as shown in Figure 11 (right), and an estimation of the background events is needed to measure correctly the mis-identification probability. After the background subtraction, the mis-identification probability is measured in data and compared with that estimated in $Z \rightarrow ee$ Monte Carlo simulations. The data/MC correction factors are then extracted from the ratio of the two probabilities. The main source of systematic uncertainties is the background subtraction and this is estimated in a conservative way by comparing the data/MC correction factors with and without the background subtraction and taking the difference as uncertainty. Another source of systematic uncertainty comes from the event selection. The identification requirement and the energy scale of the tagging electron have been varied and the observed differences in the data/MC correction factors are also taken as a systematic uncertainty. The measurement has been performed in four pseudorapidity regions, which are defined using the $\tau_{\text{had-vis}}$ leading track direction: barrel ($|\eta_{\text{trk}}| < 1.37$), crack ($1.37 < |\eta_{\text{trk}}| < 1.52$), endcap ($1.52 < |\eta_{\text{trk}}| < 2.0$) and forward endcap ($|\eta_{\text{trk}}| > 2.0$). The estimated data/MC correction factors are found to be independent of the tightness of the $\tau_{\text{had-vis}}$ identification applied to the probe $\tau_{\text{had-vis}}$ and of the type of electron overlap removal. For this reason only correction factors for different working points of the electron discrimination are reported in Table 1.

electron BDT veto	$\lvert\eta_{\text{trk}}\rvert < 1.37$	$1.37 < \lvert\eta_{\text{trk}}\rvert < 1.52$	$1.52 < \lvert\eta_{\text{trk}}\rvert < 2.00$	$\lvert\eta_{\text{trk}}\rvert > 2.00$
loose	0.96±0.22	0.8±0.3	0.47±0.14	1.7 ±0.4
medium	1.3 ±0.5	-	0.5 ±0.4	2.8 ±1.3

Table 1: The data/MC correction factors for the efficiency of the electron discrimination applied to electrons mis-identified as $\tau_{\text{had-vis}}$ with $p_{\text{T}} > 20$ GeV. The correction factors are not dependent on the tightness of the $\tau_{\text{had-vis}}$ identification or on the type of electron overlap removal. The quoted uncertainties are the sum in quadrature of statistical and systematic uncertainties. Some measurements are not available due to lack of sufficient data statistics.

5 Measurements of the $\tau_{\text{had-vis}}$ Identification Efficiency

This section describes the measurements of the $\tau_{\text{had-vis}}$ identification efficiency with 2011 data and the evaluation of the data/MC correction factors. Three measurements have been performed, two based on $Z \to \tau\tau$ and one based on $W \to \tau\nu$ events. These two processes have been chosen because they have specific signatures which can be selected by tag-and-probe methods and with such methods it is possible to get highly-pure samples of real τ_{had} without any identification requirement. The tags used to select $Z \to \tau_\mu\tau_{\text{had}}$ and $W \to \tau_{\text{had}}\nu_\tau$ events are an isolated muon and the presence of high missing transverse energy, respectively. Each measurement is different from the others not only in the event selection, but also in the background estimation strategy. Two measurements estimate the fraction of real and mis-identified $\tau_{\text{had-vis}}$ from the fit of the track multiplicity distribution of the probe $\tau_{\text{had-vis}}$ and one from a two-dimensional side-band extrapolation. Because of these differences, the definitions of the probe candidate used for the efficiency measurement are not the same. Table 2 outlines the main features of each measurement and each probe definition that is described in this section. Even though the measured efficiencies cannot be directly compared, the agreement or disagreement between what is measured in data and what is estimated in Monte Carlo simulations should be similar across all measurements.

	Section 5.2	Section 5.3	Section 5.4
Events	$Z \to \tau_\mu\tau_{\text{had}}$	$Z \to \tau_\mu\tau_{\text{had}}$	$W \to \tau_{\text{had}}\nu_\tau$
Background Estimation	side-band extrapolation	track fit	track fit
Probe Definition	$p_{\text{T}} > 20$ GeV, $\lvert\eta\rvert < 2.47$	$p_{\text{T}} > 20$ GeV, $\lvert\eta\rvert < 2.47$	$p_{\text{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$
	N_{tracks} =1 or 3	$N_{\text{tracks}} \geq 1$	$N_{\text{tracks}} \geq 1$
	|charge| = 1	|charge| = 1	-
	elec. BDT score>0.3	elec. BDT score>0.3	*loose* elec. BDT
	jet BDT score>0.3	jet BDT score>0.3	-
	$N_{\text{iso tracks}}$ <2	-	-
1-prong, 3-prong and multi-prong definitions based on:	default N_{tracks}	p_{T}-corr. N_{tracks}	p_{T}-corr. N_{tracks}

Table 2: Outline of the three $\tau_{\text{had-vis}}$ identification efficiency measurements presented in this section and definitions of each selection of the probe candidates.

The identification efficiency, ε^{id}, is defined as the fraction of the probe $\tau_{\text{had-vis}}$ (Table 2) that pass the $\tau_{\text{had-vis}}$ identification. This efficiency is measured in data and estimated in Monte Carlo simulations using the same event selection. The data/MC correction factors are extracted as:

$$C^{\text{id}}_{\text{data/MC}} = \frac{\varepsilon^{\text{id}}_{\text{Data}}}{\varepsilon^{\text{id}}_{\text{MC}}} \tag{1}$$

where $\varepsilon^{\text{id}}_{\text{Data}}$ and $\varepsilon^{\text{id}}_{\text{MC}}$ are the identification efficiencies measured in data and estimated in Monte Carlo simulations, respectively. The correction factors have been measured for different combinations of $\tau_{\text{had-vis}}$ identification, electron veto and muon veto working points, as shown in Table 3.

The $Z \to \tau\tau$ measurement using the side-band background extrapolation, Section 5.2, and the $W \to \tau\nu$ measurement, Section 5.4, were already performed with data collected in early 2011 and documented

$\tau_{\text{had-vis}}$ ID	Electron Veto	Muon Veto
BDT *loose*	BDT *loose*	no
BDT *medium*	BDT *medium*	yes
BDT *tight*	BDT *tight*	yes
LLH *loose*	BDT *loose*	no
LLH *medium*	BDT *medium*	yes
LLH *tight*	BDT *tight*	yes

Table 3: Definitions of the identification working points used to measure the data/MC correction factors.

in Ref. [15]. This section describes the update of these measurements also including the data recorded in late 2011, which are characterised by a higher average number of interactions per bunch crossing. The $Z \to \tau\tau$ measurement using the track multiplicity distribution fit, Section 5.3, is new. In Section 5.5, a summary of the results of these studies is presented. All measurements use a p_{T}-correlated track counting as track isolation requirement or in the track multiplicity distribution fit. This counting is described in Section 5.1. The measurements are performed on two different sets of events. The $Z \to \tau\tau$ analyses use data collected with a single muon trigger corresponding to an integrated luminosity of $3.6\,\text{fb}^{-1}$, while the $W \to \tau\nu$ measurement is based on data collected by missing transverse energy and missing transverse energy combined with jet triggers corresponding to an integrated luminosity of $4.6\,\text{fb}^{-1}$.

5.1 The p_{T}-correlated Track Counting

This track counting is aimed at improving the discrimination power of the track multiplicity distribution between real $\tau_{\text{had-vis}}$ and QCD jets. In the $\tau_{\text{had-vis}}$ reconstruction, the track association is performed in the core cone of radius $\Delta R \leq 0.2$ with respect to the $\tau_{\text{had-vis}}$ direction. This choice of radius is intended to ensure that only the tracks produced by the hadronic decay are selected. This means that the size of the cone is optimal for the $\tau_{\text{had-vis}}$ reconstruction, but it is unlikely that many tracks from a QCD jet can be selected in such a narrow cone. Therefore the track counting is performed in a wider cone of radius $\Delta R \leq 0.6$ to improve the discrimination power. With a cone of such a size the real $\tau_{\text{had-vis}}$ still have a track multiplicity distribution peaked at 1 and 3 and the QCD jets are characterised by a much higher number of tracks. This leads to very different track multiplicity templates to describe real and fake $\tau_{\text{had-vis}}$.

The track counting algorithm is similar to the one used in the anti-k_T- or k_T-clustering jet algorithm [20, 28]. It selects only the tracks in the region $0.2 < \Delta R \leq 0.6$ p_{T}-correlated to the ones in the *core* cone ($\Delta R \leq 0.2$) in order to avoid tracks from underlying or pile up events. The track multiplicity is then the sum of the tracks in $\Delta R \leq 0.2$ and the selected tracks in $0.2 < \Delta R \leq 0.6$. The method was first demonstrated in Ref. [1] and then used in the previous $\tau_{\text{had-vis}}$ identification efficiency measurements [15, 29]. In this note, this counting is used in Section 5.2 to apply a track isolation cut, while in Sections 5.3 and 5.4 it is used as the input variable for the track multiplicity distribution fit and defines the 1-prong and multi-prong $\tau_{\text{had-vis}}$.

5.2 Measurement of the $\tau_{\text{had-vis}}$ Identification Efficiency with $Z \to \tau\tau$ Events

This section describes the measurement of the $\tau_{\text{had-vis}}$ identification efficiency with $Z \to \tau_\mu\tau_{\text{had}}$ events in $3.6\,\text{fb}^{-1}$ of data collected in 2011. The dataset is split into two sub-samples: one corresponding to $1.8\,\text{fb}^{-1}$ collected in early to mid 2011 with low pile up (mean average number of interactions per bunch

crossing of $< \mu >= 6$) and the other corresponding to $1.9\,\text{fb}^{-1}$ collected in late 2011 with high pile up ($< \mu >= 11$). The efficiencies measured in both data periods are found to be in good agreement with the corresponding efficiency from the simulation. The analysis is performed as in Ref. [15], with additional event selection requirements (described in Section 5.2.1), which allow a higher signal purity to be obtained. The dominant backgrounds are from QCD multi-jet processes and W + jets production. The QCD multi-jet contribution is estimated from data using a two-dimensional side-band extrapolation method. The W + jets contribution is estimated from Monte Carlo simulations and scaled by a data/MC scale factor (k_W) obtained in a control region dominated by W + jets events. The remaining backgrounds where the probe $\tau_{\text{had-vis}}$ is an incorrectly selected jet or electron, like Z + jets and $t\bar{t}$ events, are estimated from simulation. The dominant systematic uncertainty comes from the normalisation of the QCD multi-jet background.

5.2.1 Event Selection

The event selection is very similar to what has been used in the previous study [15]. In the tag selection only two changes have been made. Firstly, due to the increased instantaneous luminosity in late 2011 data, the muon trigger has been tightened to keep the output event rate within the allowed bandwidth. Secondly, the isolation requirement has been tightened to increase the suppression of QCD multi-jet events, but only in the first data period. Indeed, in the second data period the average energy deposited in the calorimeter around the muon direction is increased due to the higher pile up environment and this tighter isolation cut causes a significant signal loss. The tag selection is then a muon which has no additional associated tracks in a cone of radius $\Delta R < 0.4$, and whose sum of transverse energy deposited in the annulus $0.05 < \Delta R < 0.2$ with respect to its direction is less than 2% of its transverse momentum in the first data period, or less than 4% in the second data period. The probe selection is also very similar to the previous analysis. The $\tau_{\text{had-vis}}$ candidate is required to have $p_{\text{T}} > 20$ GeV, $|\eta| < 2.47$, a charge magnitude of one and one or three associated tracks. Then, instead of applying the tight cut-base discrimination for mis-identified electrons as in the previous study, the $\tau_{\text{had-vis}}$ candidate needs to have an electron BDT score greater than 0.3. This is a looser requirement that reduces the bias on the $\tau_{\text{had-vis}}$ candidate and provides enough discrimination against backgrounds with mis-identified electrons, like $Z \to \tau_\mu \tau_e$. Figure 12 (left) shows the electron BDT score for the probe $\tau_{\text{had-vis}}$. A loose threshold of 0.3 on the jet BDT score is also applied. This requirement does not bias the probe selection, but suppresses a large amount of background. The jet BDT score is shown in Figure 12 (right). Finally, the $\tau_{\text{had-vis}}$ candidate is required to have less than two tracks in the annulus $0.2 < \Delta R < 0.4$ around its direction, counted using the algorithm described in Section 5.1. Figure 12 (bottom) shows this track multiplicity distribution. The complete probe selection is summarised in Table 2. The other event selection criteria are the same as in Ref. [15]. Events must have a reconstructed primary vertex with at least 4 associated tracks, pass a set of jet cleaning requirements, have exactly one muon, no selected electrons and at least one selected $\tau_{\text{had-vis}}$ candidate. The removal of overlapping muons, electrons and $\tau_{\text{had-vis}}$ candidates is performed. A further selection is applied to suppress the W + jets background, by requiring $\sum \cos\Delta\phi = \cos\Delta\phi(\tau_{\text{had-vis}}, E_{\text{T}}^{\text{miss}}) + \cos\Delta\phi(\mu, E_{\text{T}}^{\text{miss}}) > -0.15$ and transverse mass $m_{\text{T}} = \sqrt{2p_{\text{T}}^{\mu} \cdot E_{\text{T}}^{\text{miss}}(1 - \cos\Delta\phi(\mu, E_{\text{T}}^{\text{miss}}))} < 50$ GeV. If more than one $\tau_{\text{had-vis}}$ candidate satisfies the selection and overlap removal, the candidate with highest transverse momentum is chosen as the probe $\tau_{\text{had-vis}}$.

5.2.2 Background Estimation and Systematic Uncertainties

The contributions from the dominant processes are summarised in Table 4. The dominant background processes are W + jets and QCD multi-jet production. The QCD multi-jet contribution is estimated from data using a two-dimensional side-band extrapolation method. The side-bands are defined using

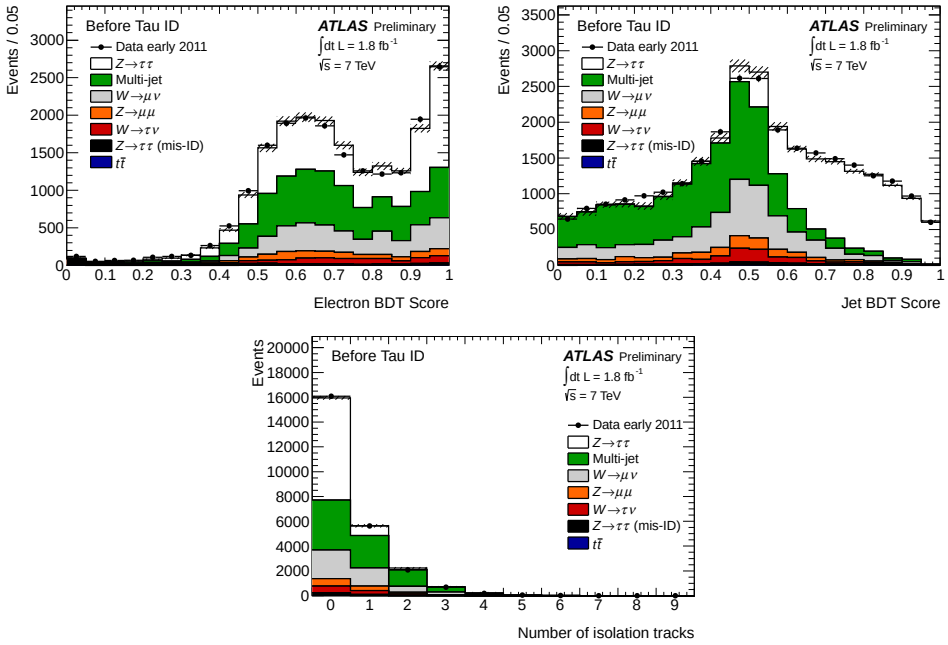


Figure 12: $Z \rightarrow \tau\tau$ (side-band extrap.) measurement: Discriminants used to increase the purity of the $\tau_{\text{had-vis}}$ sample without biasing the probe $\tau_{\text{had-vis}}$ selection. Electron and jet BDT scores are shown in the top left and right. The number of tracks in the annulus $0.2 < \Delta R < 0.4$ around the probe $\tau_{\text{had-vis}}$ based on the p_{T}-correlated counting is shown at the bottom. Each distribution is shown using only data collected in the low pile up period after the full event selection, without the cut on the variable itself.

the charge product between the tag muon and the probe $\tau_{\text{had-vis}}$, and the isolation of the tag muon. No alterations to the method have been made with respect to Ref. [15]. The W + jets contribution is taken from Monte Carlo simulations and scaled by a data/MC scale factor (k_W), measured in a control region obtained by reversing the selection on m_{T} and $\sum \cos\Delta\phi$. Different scale factors are used before and after applying the $\tau_{\text{had-vis}}$ identification. As an addition to the previous analysis, the k_W factors are measured independently for opposite-sign (OS) and same-sign (SS) events, which have been shown to give significantly different values [7]. Typical values before applying the $\tau_{\text{had-vis}}$ identification are $k_W^{\text{OS}} = 0.748 \pm 0.006$ and $k_W^{\text{SS}} = 0.812 \pm 0.011$, and after applying BDT *tight* identification are $k_W^{\text{OS}} = 0.531 \pm 0.025$ and $k_W^{\text{SS}} = 0.715 \pm 0.076$. The uncertainties include only the statistical error. These k_W factors are independent of the $\tau_{\text{had-vis}}$ transverse momentum. All other backgrounds are estimated from Monte Carlo simulations. The impact of systematic uncertainties on the muon p_{T} and reconstruction efficiency, the $\tau_{\text{had-vis}}$ energy scale and the luminosity are taken from the previous analysis. Although this is slightly conservative, as the uncertainties have been reduced [24], the experimental systematic uncertainties are small compared to the QCD multi-jet uncertainty.

As Table 4 shows, with the new event selection a purity of 42% or higher is achieved before identification for the inclusive and 1-prong measurements. The QCD multi-jets and W + jets are the dominant backgrounds, with other processes contributing about 6%. The purity for 3-prong $\tau_{\text{had-vis}}$ candidates before identification is much lower than the one for 1-prong $\tau_{\text{had-vis}}$ candidate and this leads to much larger systematic uncertainties.

	Process contributions before ID (%)		
	Inclusive	1-prong	3-prong
Signal	42	49	28
W+jets	21	19	25
QCD multi-jets	30	26	40
Other	6	6	7

Table 4: $Z \to \tau\tau$ (side-band extrap.) measurement: Fractional contributions of processes in the analysis before the $\tau_{\text{had-vis}}$ identification is applied. After applying the $\tau_{\text{had-vis}}$ identification, the purity increases to 70-92%. The purity for 3-prong candidates is lower and leads to larger systematic uncertainties.

5.2.3 The p_{T}-inclusive Measurement

This section presents the efficiencies measured in data and the corresponding data/MC correction factors for the p_{T}-inclusive measurement. The visible mass distributions of the muon-$\tau_{\text{had-vis}}$ pair system are shown in Figure 13 before and after identification. Reasonable agreement is observed between data and the sum of the estimated contributions. Since the uncertainty on the inclusive measurement is dominated by the uncertainty on the background estimation, only the data collected in the low pile up period have been used. Because of the looser isolation cut in the high pile up period, the selected events in the second data period have higher contamination from QCD multi-jet events and therefore the method is affected by higher systematic uncertainties. A summary of the efficiencies measured in data and from Monte Carlo simulations, including all measurement uncertainties is given in Table 5. The individual contributions to the uncertainty on the efficiency measured from data are: the statistical uncertainty, $\Delta\varepsilon_{\text{stat}}$, the normalisation uncertainties on the W + jets and QCD multi-jet backgrounds, $\Delta\varepsilon_{\text{W+jets}}$ and $\Delta\varepsilon_{\text{QCD}}$, and the experimental uncertainties on muons and $\tau_{\text{had-vis}}$ and on the integrated luminosity, $\Delta\varepsilon_{\text{exp}}$. The total

uncertainty is obtained from pseudo-experiments treating all uncertainties as uncorrelated. For both the 1-prong and 3-prong measurements the statistical uncertainty is comparable to the uncertainty from the QCD multi-jet normalisation, while for the combined measurement the QCD multi-jet uncertainty dominates. The data/MC correction factors for inclusive, 1-prong and 3-prong $\tau_{\text{had-vis}}$ candidates for the low pile up period are found to be consistent with unity, as summarised in Table 6. The correction factors measured in the high pile up period are also consistent with unity.

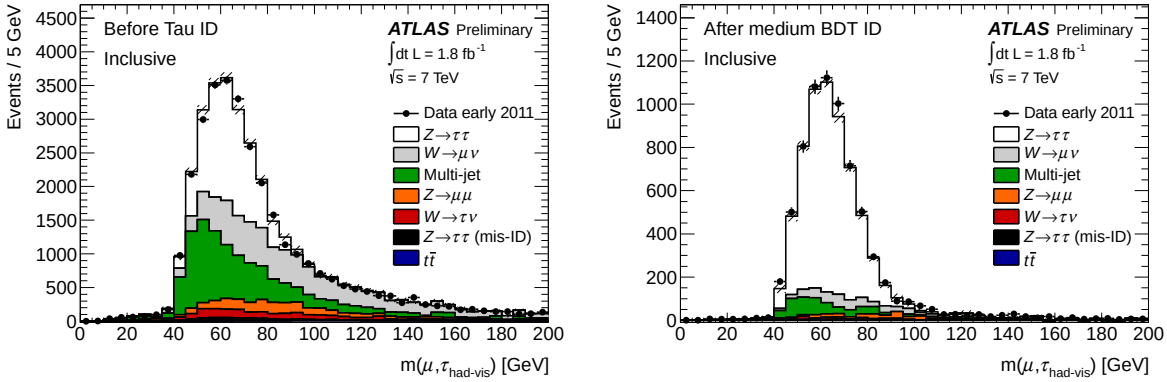


Figure 13: $Z \to \tau\tau$ (side-band extrap.) measurement: Visible mass distributions before applying $\tau_{\text{had-vis}}$ ID (left) and after applying the BDT *medium* identification (right) to the probe $\tau_{\text{had-vis}}$ for the low pile up period.

			Uncertainty contributions (%)				
ID	$\varepsilon_{\text{MC}}(\pm\text{stat})$	$\varepsilon_{\text{Data}}$	$\Delta\varepsilon_{\text{stat}}$	$\Delta\varepsilon_{\text{W+jets}}$	$\Delta\varepsilon_{\text{QCD}}$	$\Delta\varepsilon_{\text{exp.}}$	$\Delta\varepsilon_{\text{Total}}$
BDT *loose*	0.748±0.003	0.822	2.3	0.3	3.9	2.2	5.1
BDT *medium*	0.534±0.003	0.574	2.5	0.3	4.2	2.2	5.4
BDT *tight*	0.282±0.003	0.297	2.9	0.3	4.3	2.2	5.8
LLH *loose*	0.833±0.002	0.936	2.0	0.3	3.3	2.2	4.5
LLH *medium*	0.607±0.003	0.669	2.3	0.3	3.9	2.2	5.1
LLH *tight*	0.332±0.003	0.358	2.8	0.3	4.3	2.2	5.6

Table 5: $Z \to \tau\tau$ (side-band extrap.) measurement: Inclusive $\tau_{\text{had-vis}}$ identification efficiencies in MC and measured from data for the low pile up period with all measurement uncertainties. The total uncertainty is obtained from pseudo-experiments treating all uncertainties as uncorrelated.

5.2.4 The p_{T}-binned Measurement

In this section, the study is repeated in bins of the p_{T} of the probe $\tau_{\text{had-vis}}$ candidate. The method is identical, except that a sliding window on the visible mass is used to increase the signal purity in each $\tau_{\text{had-vis}}$ p_{T} bin: $x < m(\mu, \tau_{\text{had-vis}}) < x + 25$ GeV, where x is the sum of the lower thresholds on the tag muon and probe $\tau_{\text{had-vis}}$. The threshold on the muons is 20 GeV and the $\tau_{\text{had-vis}}$ p_{T} bins are: 20, 25, 30, 35, 40, 50, 60 GeV. To increase the sample size in each bin, the low and high pile up periods are combined together. Figure 14 shows the $\tau_{\text{had-vis}}$ identification efficiencies measured in data and estimated in Monte Carlo simulation for 1-prong and 3-prong candidates for the BDT *medium* identification working point.

ID	inclusive	1-prong	3-prong
BDT *loose*	1.10±0.06	1.07±0.04	1.18±0.13
BDT *medium*	1.07±0.06	1.05±0.05	1.16±0.13
BDT *tight*	1.05±0.06	1.00±0.05	1.19±0.14
LLH *loose*	1.12±0.05	1.09±0.04	1.23±0.11
LLH *medium*	1.10±0.06	1.06±0.05	1.23±0.13
LLH *tight*	1.08±0.06	1.04±0.05	1.19±0.14

Table 6: $Z \to \tau\tau$ (side-band extrap.) measurement: p_{T}-inclusive data/MC correction factors including the combined systematic and statistical uncertainty measured in the low pile up period.

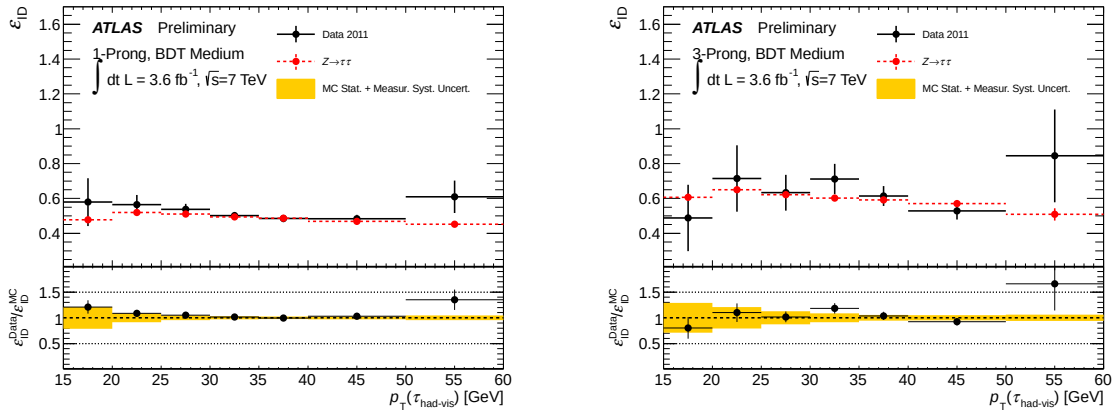


Figure 14: $Z \to \tau\tau$ (side-band extrap.) measurement: $\tau_{\text{had-vis}}$ identification efficiencies in bins of the $\tau_{\text{had-vis}}$ p_{T} for 1-prong (left) and 3-prong (right) candidates for the BDT *medium* identification working point. The errors on the measured efficiencies include systematic and statistical uncertainties, while the errors on the simulated efficiencies are only statistical uncertainties. The data/MC correction factors are shown at the bottom and their error bars include only the statistical uncertainty. The yellow band includes the systematic uncertainty of the measurement and the statistical uncertainty of the simulated efficiencies.

Table 7 summarises the data/MC correction factors, SF, for the BDT *medium* working point for the inclusive case and details each of the individual contributions to the uncertainty: the statistical uncertainty, ΔSF_{stat}, the normalisation uncertainties on the W + jets and QCD multi-jet backgrounds, ΔSF_{W+jets} and ΔSF_{QCD}, and the experimental uncertainties on the muon, $\tau_{\text{had-vis}}$ and the integrated luminosity, ΔSF_{exp}. The uncertainties for the other working points are similar. For 1-prong (3-prong) candidates, the systematic uncertainty from the QCD multi-jet normalisation is slightly smaller (larger), due to the lower (higher) level of QCD multi-jet contamination. As Figure 14 shows, at low p_T the measurement is dominated by the systematic uncertainty on the QCD multi-jet normalisation, due to the higher level of QCD multi-jet contamination. In the mid-range the uncertainties are very low, and the efficiency ratios are in good agreement with unity. At high-p_T the measurement is dominated by statistical uncertainty as there are very few $Z \to \tau\tau$ events in this range. For most working points, the ratios are in agreement with unity within uncertainty over the whole p_T range.

		data/MC correction factor uncertainty contributions (%)				
p_T [GeV]	SF	ΔSF_{stat}	ΔSF_{W+jets}	ΔSF_{QCD}	ΔSF_{exp}	ΔSF_{Total}
15–20	1.14 ± 0.30	9.8	1.5	21.6	1.3	26.6
20–25	1.08 ± 0.13	5.4	0.7	10.2	1.4	12.0
25–30	1.04 ± 0.07	3.8	0.5	5.4	1.4	7.0
30–35	1.05 ± 0.05	3.3	0.4	3.4	1.4	5.2
35–40	1.01 ± 0.05	3.3	0.4	2.3	1.5	4.6
40–50	0.99 ± 0.05	3.6	0.4	2.1	1.5	4.8
50–60	1.49 ± 0.29	17.9	2.3	4.9	1.0	19.5

Table 7: $Z \to \tau\tau$ (side-band extrap.) measurement: Inclusive data/MC correction factors for the BDT *medium* working point, measured in bins of $\tau_{\text{had-vis}}$ p_T with statistical and systematic uncertainties.

5.3 Alternative Method for Measurement of the $\tau_{\text{had-vis}}$ Identification Efficiency with $Z \to \tau\tau$ Events

This section describes the measurement of the $\tau_{\text{had-vis}}$ identification efficiency with the same $Z \to \tau_\mu \tau_{\text{had}}$ events as the ones used in Section 5.2, but with an alternative background estimation based on the fit to the track multiplicity spectrum of the $\tau_{\text{had-vis}}$ candidate, similar to that used also in Section 5.4. The track multiplicity used is based on the p_T-correlated counting described in Section 5.1 and offers a good separation between signal and background. This background estimation does not suffer the limitation coming from background normalisation which affects the method described in the previous section. In this section, if not mentioned otherwise, the track multiplicity of the probe $\tau_{\text{had-vis}}$ is computed with the p_T-correlated algorithm.

5.3.1 Event Selection

The selection applied on the tagging muon is similar to that used in Section 5.2.1 with some differences. For the entire dataset, the sum of transverse energy deposited in the annulus $0.05 < \Delta R < 0.2$ with respect to the muon direction has to be less than 4% of the muon transverse momentum. In addition, $\Delta\phi(\tau_{\text{had-vis}}, \mu) > 2.4$ and $p_T^\mu < 40$ GeV are required to suppress the W + jets background. This different event selection is motivated by the fact that the requirement of one or three tracks used in Section 5.2.1

cannot be applied here. Indeed, since the background estimation is based on the fit of the track multiplicity distribution of the probe $\tau_{\text{had-vis}}$, all candidates with at least one track are accepted. Such a looser selection decreases the purity of the $\tau_{\text{had-vis}}$ candidate before applying the $\tau_{\text{had-vis}}$ identification and, therefore, a tighter event selection is needed. Such differences in the selection of the probe $\tau_{\text{had-vis}}$ have to be taken into account when comparing the $\tau_{\text{had-vis}}$ identification efficiencies and the data/MC correction factors measured with this background estimation with the ones in Section 5.2. The estimated contributions before identification are reported in Table 8, for the signal region, where opposite-signed charges between the probe $\tau_{\text{had-vis}}$ and the muon are required (OS), and for the same-sign region, where same-signed charges are required (SS). The W + jets background is scaled by the k_W factors described in Section 5.2.2.

Process contributions before ID (%)		
	OS	SS
Signal	32	1
W+jets	18	17
QCD multi-jets	46	76
Other	4	6

Table 8: $Z \to \tau\tau$ (track fit) measurement: Fractional contributions of processes before $\tau_{\text{had-vis}}$ identification. The W + jets background is scaled by the k_W factors described in Section 5.2.2. Values are shown in the signal region (OS) and in the control region (SS). After the BDT *medium* identification a signal purity of 78% is reached.

The background is dominated by QCD multi-jet and W+jets events. In the fit only two templates are used: the signal template, which is taken from $Z \to \tau\tau$ Alpgen MC simulation requiring truth-matching in the OS signal region, and the background template, which is taken from data in the SS control region. The background template models the QCD jets from both QCD multi-jet and W + jet events. No electron template is needed, given that the electron contribution in the track multiplicity spectrum is strongly suppressed by selecting candidates with an electron BDT score > 0.3. In general no separate lepton template is needed, because the electron contribution is negligible, and the Z$\to \mu\mu$ background is already included in the total background template.

The W + jets background is known to have a charge correlation and this causes the OS region to be more enriched in quark-initiated jets, that are then mis-identified as $\tau_{\text{had-vis}}$ candidates, with respect to the SS region. In general, quark- and gluon-initiated jets have different track multiplicity distributions. However the difference is minimised by applying a loose cut on the jet BDT score, as shown in Figure 15, where the track multiplicity distribution for simulated $W \to \mu\nu$+jets events is plotted in OS and SS regions, before and after applying this cut. The drawback of this is that the background template becomes more peaked at lower values of the track multiplicity, but it still provides adequate discrimination.

5.3.2 Fitting Strategy

The efficiency of all ID methods and working points is measured using a single simultaneous fit to the track multiplicity distributions in data, obtained before and after applying each of the different identification working points. In order to measure the 1-prong and multi-prong efficiency, the fit is performed splitting the signal template in two parts: the first bin, for the measurement of the 1-prong efficiency, and the rest of the spectrum for the measurement of the multi-prong efficiency. For the measurement of the inclusive efficiencies, the signal template is not split. As an example of the performance of the fit, Fig-

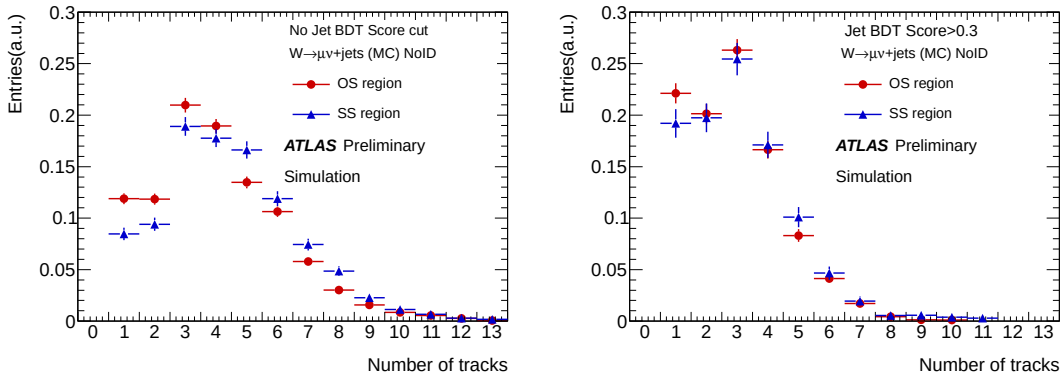


Figure 15: $Z \to \tau\tau$ (track fit) measurement: Track multiplicity distributions for simulated $W \to \mu\nu$+jets background in OS and SS regions before $\tau_{\text{had-vis}}$ ID. The left plot shows the distributions with no requirement on the jet BDT score, and the right plot shows the distributions selecting candidates with a jet BDT score > 0.3. The application of this selection reduces the differences in the shapes between OS and SS regions.

ure 16 shows the result for the inclusive measurement before identification (top left) and after the BDT *loose* (top right), the *medium* (bottom left) and the *tight* (bottom right) identification working points. Pseudo-experiments are used to check the stability of the fitting method against correlations between the templates and it is found that no corrections are needed.

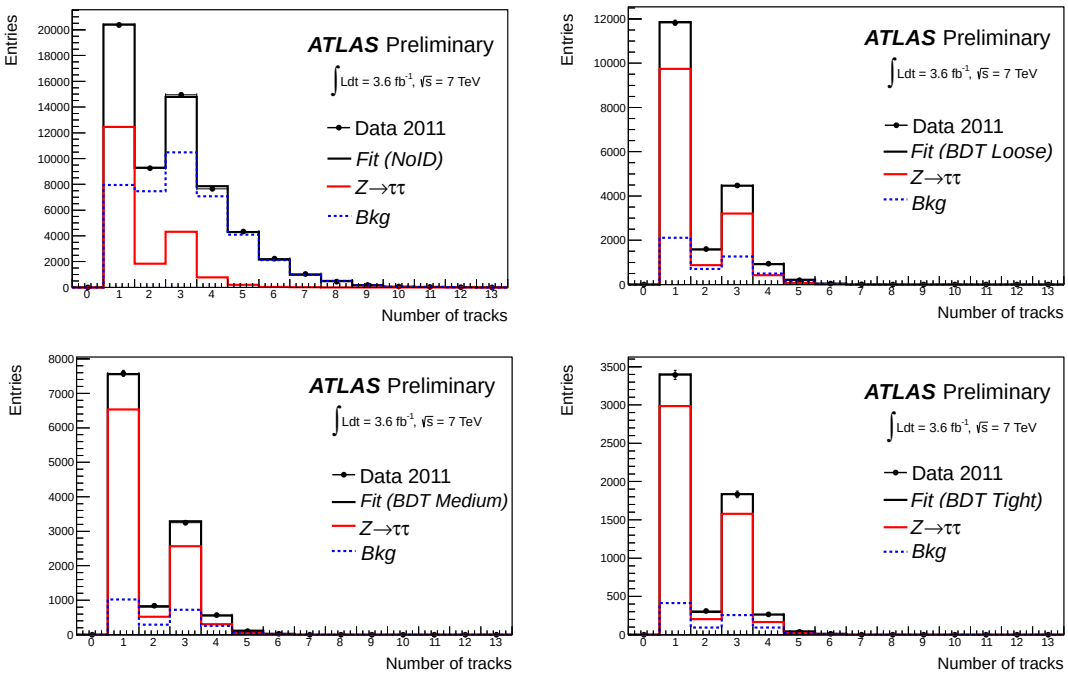


Figure 16: $Z \to \tau\tau$ (track fit) measurement: Simultaneous fit for the measurement of the inclusive identification efficiencies. Here only the BDT identification working points are shown, but in the fit all identification methods and all working points are included at the same time.

5.3.3 Systematic Uncertainties

The effect of various sources of systematic errors has been estimated for both signal and background.

Signal template: Effects that could change the signal model are studied by repeating the fit with signal templates generated from Monte Carlo simulations with different configurations. These new signal templates are obtained from different Monte Carlo generators (ALPGEN vs. PYTHIA), different detector geometries, different hadronic shower models (QGSP [30] and FTFP_BERT [31]), and different Monte Carlo tuning (Perugia [32]) since they all might affect the track multiplicity distributions before and after $\tau_{\text{had-vis}}$ identification.

Pile up: The effect of different pile up conditions has been estimated for signal by repeating the fit with two different signal templates built in low pile up conditions (number of additional vertices ≤ 6) and in high pile up conditions (number of additional vertices > 6).

Background template: Differences in the background shapes between the OS signal region and SS control region (OS-SS uncertainty) mainly come from the different relative fractions of quark- and gluon-initiated jets in the two regions. This is correlated to the different contaminations of $W+$jets in the two regions. In order to estimate this uncertainty, the fit is repeated after subtracting the $W \to \mu\nu$+jets contribution in SS region from the background template (after properly scaling this contribution with the scale factors described in Section 5.2.2) and adding the $W \to \mu\nu$+jets contribution in the OS region.

The full set of systematic uncertainties for the inclusive measurements is listed in the Table 9. For the inclusive, as well as for the 1-prong, measurement the dominant systematic uncertainty is represented by the hadronic shower model. This is due to the lack of statistics in the simulation sample used for the comparison and not to the method itself. The uncertainties for the multi-prong measurements are bigger and dominated by the OS-SS uncertainty.

$p_{\text{T}} > 20$ GeV	inclusive efficiency measurement systematic uncertainties (%)						
ID	Generator	Hadronic Shower	Geometry	MC Tune	Pile up	OS-SS	Total
BDT *loose*	0.1	1.8	2.6	1.9	0.5	1.4	4.1
BDT *medium*	0.3	4.8	0.3	3.6	0.8	2.1	6.4
BDT *tight*	0.4	16.6	5.3	2.6	2.4	0.8	18.1
LLH *loose*	0.5	2.0	2.5	0.25	1.2	1.4	3.8
LLH *medium*	0.8	0.9	0.9	2.5	3.1	1.5	4.5
LLH *tight*	0.1	5.5	2.1	3.3	3.1	2.6	7.9

Table 9: $Z \to \tau\tau$ (track fit) measurement: Systematic uncertainties on the inclusive efficiency measurement. The dominant uncertainty is represented by the shower shape modeling. The total uncertainty is the quadratic combination of all the listed uncertainties.

5.3.4 The p_{T}-inclusive Measurement

The efficiencies measured in data with statistical and systematic uncertainties for the inclusive case are listed in Table 10. The efficiencies expected from Monte Carlo simulations are also quoted along with their statistical uncertainty. The data/MC correction factors are listed in Table 11.

			Uncertainty contributions (%)		
ID	$\varepsilon_{\mathrm{MC}}$	$\varepsilon_{\mathrm{Data}}$	$\Delta\varepsilon^{stat}_{\mathrm{Fit}}$	$\Delta\varepsilon^{syst}_{\mathrm{Fit}}$	$\Delta\varepsilon_{\mathrm{Total}}$
BDT *loose*	0.717±0.008	0.729	2.1	4.1	4.7
BDT *medium*	0.508±0.006	0.508	2.3	6.4	6.9
BDT *tight*	0.263±0.004	0.313	3.1	18.1	18.4
LLH *loose*	0.809±0.009	0.826	1.9	3.8	4.4
LLH *medium*	0.587±0.006	0.589	2.1	4.5	5.1
LLH *tight*	0.318±0.004	0.313	2.5	7.8	8.4

Table 10: $Z \to \tau\tau$ (track fit) measurement: p_{T}-inclusive efficiencies in MC and measured in data with the measurement uncertainties.

ID	inclusive	1-prong	multi-prong
BDT *loose*	1.02±0.05	1.02±0.05	1.00±0.25
BDT *medium*	1.00±0.07	1.00±0.07	0.95±0.24
BDT *tight*	0.96±0.18	0.96±0.13	0.93±0.27
LLH *loose*	1.02±0.04	1.03±0.03	1.08±0.20
LLH *medium*	1.00±0.05	1.01±0.03	1.04±0.16
LLH *tight*	0.98±0.08	0.98±0.06	0.91±0.25

Table 11: $Z \to \tau\tau$ (track fit) measurement: p_{T}-inclusive data/MC correction factors including the combined systematic and statistical uncertainty.

5.3.5 The p_{T}-binned Measurement

The same measurement is also performed in bins of p_{T} of the probe $\tau_{\text{had-vis}}$. Table 12 compares the data/MC correction factors in the p_{T}-binned and p_{T}-inclusive measurements for the BDT *medium* identification working point. The measurement is totally dominated by systematic uncertainties, even in the low- and high-p_{T} bins where the statistical uncertainties are increased because of the higher background contamination and the smaller statistics, respectively.

BDT ID (inclusive)	*medium*		
p_{T} bin	SF	$\Delta\mathrm{SF}^{stat}$(%)	$\Delta\mathrm{SF}^{syst}$(%)
$p_{\mathrm{T}} > 20$ GeV	1.00	2.6	6.4
$15 < p_{\mathrm{T}} < 20$ GeV	0.85	7.7	28
$20 < p_{\mathrm{T}} < 25$ GeV	1.08	5.8	17
$25 < p_{\mathrm{T}} < 30$ GeV	1.01	5.2	18
$30 < p_{\mathrm{T}} < 40$ GeV	0.94	4.5	15
$p_{\mathrm{T}} > 40$ GeV	0.94	7.6	32

Table 12: $Z \to \tau\tau$ (track fit) measurement: Data/MC correction factors measured inclusively and in p_{T} bins for the BDT *medium* identification working point.

5.4 Measurement of the $\tau_{\text{had-vis}}$ Identification Efficiency with $W \to \tau\nu$ Events

In this section, the measurement of the $\tau_{\text{had-vis}}$ identification efficiency using $W \to \tau\nu$ events is presented. The measurement is documented in Refs. [15, 29] for the early data taking period and it is repeated here for all 2011 data. The analysis is a tag-and-probe method where the events are tagged by a large missing transverse energy and the probe is a reconstructed $\tau_{\text{had-vis}}$. Because of this tag selection, $W+ \geq 1\,\text{jet}$ events are enhanced. The number of real $\tau_{\text{had-vis}}$ before and after identification is estimated by fitting the distribution of the number of tracks associated to the candidate using the p_{T}-correlated counting, as described in Section 5.1. The fit is performed using three templates, one for real $\tau_{\text{had-vis}}$, one for mis-identified muons and electrons and one for mis-identified QCD jets. The tau and lepton templates are taken from Monte Carlo simulations, while the jet template is extracted from data in a control region. If not specified otherwise, in this section the track multiplicity is computed with the p_{T}-correlated algorithm.

5.4.1 Event Selection

The event selection is the same as the one used in Ref. [15] except for the tag selection. Due to the higher instantaneous luminosity in late 2011, the trigger threshold for missing transverse energy was raised from 60 to 80 GeV and accordingly also the offline tag selection has been tightened. Therefore, in this measurement events are selected if they are triggered by a missing transverse energy trigger or a combined trigger requiring both missing transverse energy and a jet and they have more than 100 GeV of reconstructed missing transverse energy. Moreover, in order to reject events with mis-measured missing transverse energy and QCD multi-jets events, no jets or $\tau_{\text{had-vis}}$ with $p_{\mathrm{T}} > 20$ GeV may be found within $\Delta\phi < 0.5$ with respect to the direction of the missing transverse energy. As in Ref. [15], the presence of a good vertex, event quality criteria and the lepton veto is required. Furthermore, in order to increase the purity of the selection, the probe $\tau_{\text{had-vis}}$ is required to have $\Delta\phi(\tau_{\text{had-vis}}, E_{\mathrm{T}}^{\mathrm{miss}}) < 1.5$ and

$m_T = \sqrt{2E_{\mathrm{T}}^{\mathrm{miss}} p_{\mathrm{T}}(\tau_{\text{had-vis}}) \cos\Delta\phi(\tau_{\text{had-vis}}, E_{\mathrm{T}}^{\mathrm{miss}})} < 80$ GeV. In case that more than one candidate fulfills these requirements, then only the one with m_T closest to 50 GeV is chosen. Table 13 shows the expected contributions of processes estimated with Monte Carlo simulations, expressed as fractions of the total number of events seen in the data. The contribution from QCD multi-jets events is not estimated.

Process contributions before ID (%)	
Signal	21
W + jets (jet)	23
W + jets (lep)	5

Table 13: $W \to \tau\nu$ measurement: Fractional contributions of processes in the analysis yields before the $\tau_{\text{had-vis}}$ identification estimated by Monte Carlo predictions. The contribution from QCD multi-jet events is not estimated. The label "lep" means that the selected probe $\tau_{\text{had-vis}}$ is truth matched to a lepton and the label "jet" means that the probe $\tau_{\text{had-vis}}$ is not truth matched to any τ_{had} or lepton.

5.4.2 Track Templates

In the fit of the track multiplicity distribution, three different templates are used. The ones for real $\tau_{\text{had-vis}}$ and for leptons (e/μ) are taken from truth-matched candidates in $W \to \tau\nu$ and $W \to e\nu, \mu\nu$ Monte Carlo simulation samples, respectively. The template for QCD jets is extracted from data events in a control region which is defined by the same selection as the one used for the signal region, but with at least two leptons. This control region has changed with respect to the previous study [15] where at least one lepton was required. This new template has smaller contamination from real $\tau_{\text{had-vis}}$ and it could not be used in the previous measurement because of the smaller data set. The jet templates after the $\tau_{\text{had-vis}}$ identification have been compared with Monte Carlo simulations where the selected probe is a mis-identified jet. Good agreement is observed and any remaining difference is taken as a shape uncertainty.

5.4.3 Track Fit

The track fit is performed simultaneously before identification and after the *loose*, *medium* and *tight* identification as described in Ref. [15] and in Section 5.3.2. The fit parameter for the *loose* identification efficiency is constrained to be greater than that for the *medium* efficiency, which is in turn constrained to be greater than that for the *tight* efficiency. The lepton normalisation is constrained by Monte Carlo prediction and this allows the 1-prong and multi-prong efficiencies to be measured separately. The same linear corrections estimated in Ref. [15] to correct for correlations between the three templates are applied to correct biases in the fit. These corrections are estimated with pseudo-experiments by performing a linear fit of the estimated efficiencies as a function of the efficiencies used to generate the pseudo-data. They are less than 3% in all the working points.

5.4.4 Systematic Uncertainties

The systematic uncertainties considered are on the lepton normalisation and on the shapes of the track multiplicity templates for taus and jets. The lepton normalisation before $\tau_{\text{had-vis}}$ identification is constrained to the Monte Carlo prediction. After $\tau_{\text{had-vis}}$ identification, a conservative 50% uncertainty is applied to the predicted yields to account for the uncertainties in the electron to $\tau_{\text{had-vis}}$ fake rates (see

Section 4). The shape uncertainty for the tau template is estimated comparing Monte Carlo simulations with different detector geometries, different hadronic shower models and different underlying events, similar to what is done in Section 5.3.3. The shape uncertainty for the jet template comes from the different relative fraction of quark- and gluon-initiated jets in the signal and control regions. This systematic uncertainty is estimated by varying the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold from 100 GeV down to 70 GeV where the fraction of QCD multi-jets events is enhanced. In addition, the jet template is also compared with Monte Carlo simulations since the template from the data has less statistics. The impact of different pile up conditions on the track counting is estimated by taking the difference of the nominal template with the templates built with events with either $\mu < 6$ or $\mu > 12$. This uncertainty is estimated for both the tau and the jet templates. The main source of systematic uncertainties is the hadronic shower model. Table 14 summarises all sources of systematic uncertainties estimated for the BDT identifications.

Template	Systematics (%)	Inclusive			1-prong			Multi-prong		
		loose	*medium*	*tight*	*loose*	*medium*	*tight*	*loose*	*medium*	*tight*
Lepton	Electron Normalization	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Tau	Underlying Event	2.4	0.8	1.7	1.0	0.3	0.8	5.0	1.7	3.0
	Geometry	0.0	1.0	0.6	0.4	0.1	0.7	1.5	2.9	3.3
	Hadronic Shower	4.8	3.1	2.3	2.8	2.9	1.7	8.5	3.6	3.4
Jet	quark/gluon Fraction	2.8	3.0	3.0	2.5	2.6	2.5	3.3	3.8	3.9
	Shape	4.9	2.9	1.7	2.1	2.5	1.3	9.9	3.7	2.5
Tau/Jet	Pile up	0.8	1.6	1.2	0.8	1.4	0.8	0.9	2.1	1.9
Total		7.8	5.7	4.7	4.5	4.8	3.5	14.5	7.5	7.5

Table 14: $W \to \tau\nu$ measurement: Systematic uncertainties for BDT identification for $\tau_{\text{had-vis}}$ candidates with $p_T \geq 22$ GeV. The total uncertainty is the quadrature sum of each systematic component, so it does not always represent the actual overall uncertainty from the fitter.

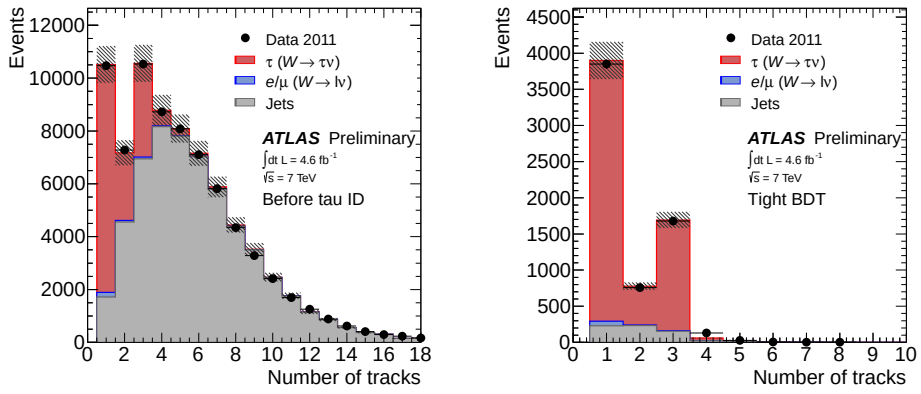
5.4.5 Results

For $\tau_{\text{had-vis}}$ with transverse momentum higher than 22 GeV, the efficiency estimated in data is consistent with the one in Monte Carlo simulations, but for $\tau_{\text{had-vis}}$ with transverse momentum lower than 22 GeV the agreement is worse, as Table 15 reports. More data are needed to understand whether this tension is really due to mis-modeling of $\tau_{\text{had-vis}}$ in the Monte Carlo simulations or due to mis-modeling of the jet template. Therefore, in this study only the results of the identification efficiency measurement for $\tau_{\text{had-vis}}$ with $p_{\mathrm{T}} > 22$ GeV are reported.

p_{T} range	BDT *loose*	BDT *medium*	BDT *tight*	LLH *loose*	LLH *medium*	LLH *tight*
$p_{\mathrm{T}} > 20$ GeV	0.93 ± 0.03	0.92 ± 0.03	0.96 ± 0.04	0.90 ± 0.04	0.93 ± 0.04	0.92 ± 0.04
$p_{\mathrm{T}} > 25$ GeV	1.00 ± 0.03	0.97 ± 0.04	1.03 ± 0.03	0.95 ± 0.05	1.01 ± 0.03	1.00 ± 0.04
$20 < p_{\mathrm{T}} < 25$ GeV	0.85 ± 0.07	0.88 ± 0.07	0.91 ± 0.09	0.84 ± 0.09	0.89 ± 0.10	0.87 ± 0.08
$p_{\mathrm{T}} > 22$ GeV	0.96 ± 0.03	0.93 ± 0.04	0.99 ± 0.04	0.95 ± 0.06	0.97 ± 0.04	0.97 ± 0.05

Table 15: $W \to \tau\nu$ measurement: Inclusive data/MC correction factors in different p_{T}-ranges. The errors include the statistical and systematic uncertainties estimated by the fitter.

The results of the track multiplicity fitting for $\tau_{\text{had-vis}}$ candidates with $p_T \geq 22$ GeV are shown in Figures 17 (a) before and (b) after the BDT *tight* identification as an example. Note that the fitting is simultaneously done for all the identification working points, i.e. for *loose*, *medium* and *tight*. The performance of this method is shown in Figure 18 where (a) the BDT score, (b) the likelihood score, (c) the calorimeter radius R_{Calo} and (d) the track radius R_{track} before identification are plotted. R_{Calo} and R_{track} are defined in Appendix A. An excellent agreement is found between data and simulation.



(a) (b)

Figure 17: $W \to \tau\nu$ measurement: Track multiplicity distributions (a) before and (b) after the BDT *tight* identification for $\tau_{\text{had-vis}}$ candidates with $p_T \geq 22$ GeV. Here only the *tight* identification is shown but the efficiency is simultaneously determined by the track multiplicity fitting in all the identification working points.

The measured identification efficiencies are summarised in Table 16 for all working points and for 1-prong plus 3-prong candidates. In this measurement, 1- and 3-prong refer to the number of tracks computed with the p_{T}-correlated counting. Therefore these efficiencies are not exactly the same as the ones for 1-prong and 3-prong $\tau_{\text{had-vis}}$ in the default definition. The agreement between data and Monte Carlo predictions is within approximately ±5% which is covered by the systematic uncertainty.

ID	$\varepsilon_{\text{Data}}$	data/MC correction factor
BDT *loose*	0.73 ± 0.03(stat) ± 0.04(syst)	0.96 ± 0.04(stat) ± 0.05(syst)
BDT *medium*	0.59 ± 0.02(stat) ± 0.03(syst)	0.93 ± 0.04(stat) ± 0.04(syst)
BDT *tight*	0.37 ± 0.01(stat) ± 0.01(syst)	0.99 ± 0.04(stat) ± 0.03(syst)
LLH *loose*	0.79 ± 0.05(stat) ± 0.04(syst)	0.93 ± 0.06(stat) ± 0.05(syst)
LLH *medium*	0.70 ± 0.03(stat) ± 0.03(syst)	0.97 ± 0.04(stat) ± 0.05(syst)
LLH *tight*	0.46 ± 0.02(stat) ± 0.03(syst)	0.96 ± 0.05(stat) ± 0.06(syst)

Table 16: $W \to \tau\nu$ measurement: Identification efficiencies and data/MC correction factors for 1- plus 3-prong $\tau_{\text{had-vis}}$ candidates with $p_T \geq 22$ GeV.

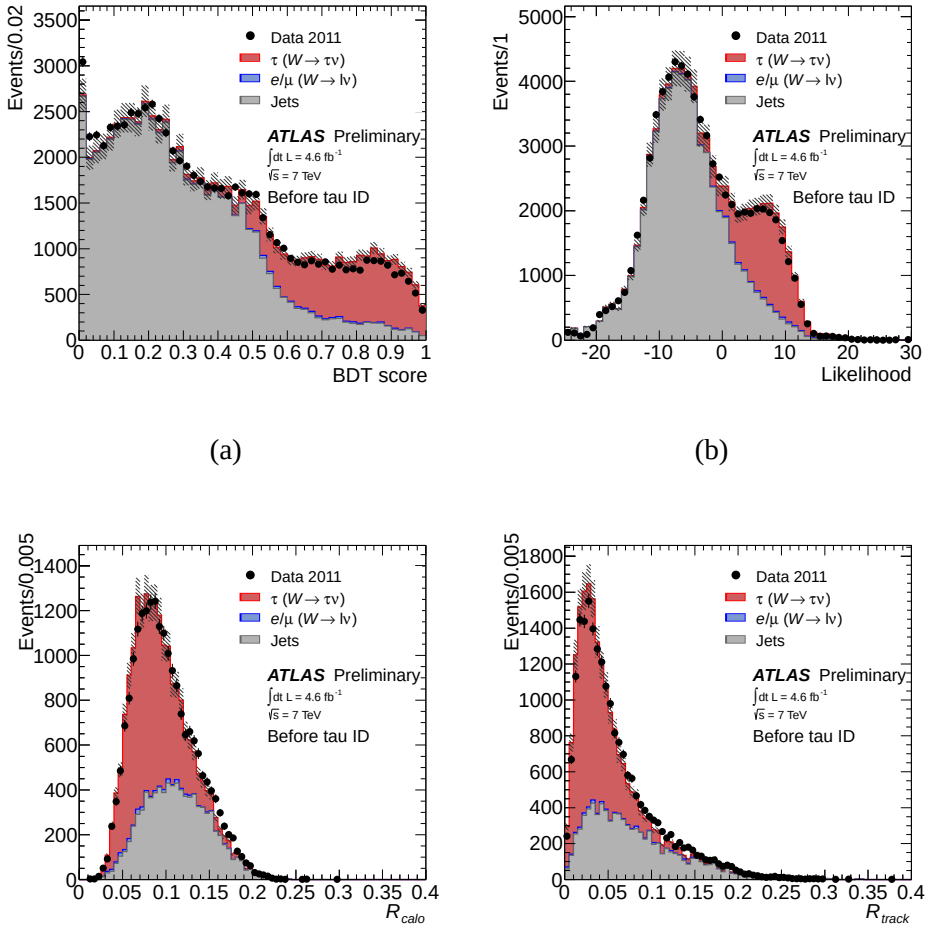


Figure 18: $W \rightarrow \tau\nu$ measurement: (a) The BDT score, (b) the likelihood score, (c) the calorimeter radius and (d) the track radius distributions before identification for $\tau_{\text{had-vis}}$ candidates with $p_T \geq$ 22 GeV. The normalisations of real $\tau_{\text{had-vis}}$, leptons and jets come from the track multiplicity fitting.

5.5 Summary of the $\tau_{\text{had-vis}}$ Identification Efficiency Measurements

Three different measurements have been presented. The first measurement (Section 5.2) is based on $Z \to \tau\tau$ events and uses a two-dimensional side-band extrapolation for the background estimation. The second one (Section 5.3) is also based on $Z \to \tau\tau$ events, but the background estimation is performed with a fit of the track multiplicity distribution of the probe $\tau_{\text{had-vis}}$. The third measurement (Section 5.4) is based on $W \to \tau\nu$ events and also uses a fit of the track multiplicity of the probe $\tau_{\text{had-vis}}$ for the background estimation. As described in Table 2, all of these measurements have many common features, but also significant differences, especially in the definition of the probe $\tau_{\text{had-vis}}$. A direct comparison of identification efficiencies and data/MC correction factors cannot be made since the definitions of the denominator in ε^{id} are different. Nonetheless it is remarkable that no significant deviation from unity has been observed in the correction factors in all three measurements. This means that the $\tau_{\text{had-vis}}$ identification efficiency is well reproduced in simulations.

Based on the results of these studies, the data/MC correction factors for 2011 data are set to unity for all the identification working points, for any track multiplicity and any transverse momentum of the $\tau_{\text{had-vis}}$ candidate. The uncertainties on these factors are estimated from the first and the third measurements presented in this section which are the most precise. Since these studies are almost orthogonal in the event selections and the background estimations, for each working point and for $\tau_{\text{had-vis}}$ candidate with $p_{\text{T}} > 22$ GeV the smallest of the two estimated uncertainties is used and listed in Table 17.

	$20 < p_{\text{T}} \leq 22$ GeV			$p_{\text{T}} > 22$ GeV		
$\tau_{\text{had-vis}}$ ID	Inclusive	1-prong	3-prong	Inclusive	1-prong	3-prong
BDT *loose*	8	8	16	4	4	8
BDT *medium*	8	10	16	4	5	8
BDT *tight*	8	8	14	4	4	7
LLH *loose*	10	8	20	5	4	10
LLH *medium*	8	10	20	4	5	10
LLH *tight*	10	10	22	5	5	11

Table 17: Uncertainty in percent on the $\tau_{\text{had-vis}}$ identification efficiency data/MC correction factors for 2011 data for different working point, different ranges of $\tau_{\text{had-vis}}$ transverse momentum and number of tracks. The data/MC correction factors are one for all identification working points, all numbers of tracks and for any transverse momentum.

In the p_{T}-range below 22 GeV, bigger uncertainties are provided in order to cover the discrepancy between the two measurements. In fact, the correction factors in this range are still compatible with one within 2σ, but a slight indication of opposite trends in the two measurements has been observed. In this range, the uncertainties in both studies are dominated by the uncertainties on the background estimation, therefore it is hard to directly compare the results. Possible biases from the different background estimations or the different probe definitions might affect the measurement in different directions.

The second measurement with $Z \to \tau\tau$ events also measured correction factors compatible with unity in all the $\tau_{\text{had-vis}}$ p_{T}-range confirming the results of the other studies. However, due to the large uncertainties, it is difficult to draw conclusions from the comparison with the other two measurements.

6 Conclusions

The re-optimised algorithms used for reconstruction and identification of τ_{had} in the full 2011 dataset were described. The re-optimisation was performed using both QCD multi-jet and $Z \to ee$ background samples selected from 2011 ATLAS data and Monte Carlo simulated signal samples. With respect to a previous implementation described in Ref. [2], the identification algorithms became less sensitive to pile up effects by developing a new algorithm to select the correct primary vertex, by increasing the use of tracking-based variables and by binning the discriminants in variables sensitive to pile up conditions. Jet rejection factors up to 130 at 50% signal efficiency and electron rejection factors of approximately 300 for 50% signal efficiency were achieved. A new procedure has been also developed to reject mis-identified muons. For a signal identification efficiency of 96%, a muon rejection factor of approximately 2 is achieved.

Measurements of the $\tau_{\text{had-vis}}$ identification efficiency and the electron to $\tau_{\text{had-vis}}$ mis-identification probabilities were made using $Z \to \tau\tau$, $W \to \tau\nu$ and $Z \to ee$ tag-and-probe analyses on 2011 data. The electron to $\tau_{\text{had-vis}}$ mis-identification probability is found to be independent of the $\tau_{\text{had-vis}}$ identification working point and of the electron overlap removal. Data/MC correction factors for the mis-identification probabilities are estimated for different electron BDT working points and for different ranges in pseudorapidity. The $\tau_{\text{had-vis}}$ identification efficiency has been measured in three different analyses based on $Z \to \tau\tau$ and $W \to \tau\nu$ events. The measured efficiencies are in agreement with the ones estimated in Monte Carlo simulations and the data/MC correction factors are consistent with unity for all the identification working points. The uncertainties on these correction factors are dominated by systematic uncertainties and range from 4% to 5% for $\tau_{\text{had-vis}}$ with $p_{\text{T}} > 22$ GeV. Larger uncertainties affect the p_T range from 20 to 22 GeV. A new method to measure the identification efficiency in the $Z \to \tau\tau$ events has been presented. This study uses a background estimation similar to the one used in the $W \to \tau\nu$ tag-and-probe analysis. Results from this study confirm what is measured by the other two analyses. Studies are ongoing to try to understand the correlations between these measurements and to improve the background estimation strategies in order to reduce the uncertainties, especially in the low-p_{T} range.

References

[1] The ATLAS Collaboration, *Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics*, CERN-OPEN-2008-020, (2008) , ISBN:978-92-9083-321-5.

[2] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*, ATLAS-CONF-2011-077, (2011). [http://cdsweb.cern.ch/record/1353226].

[3] K. Nakamura et al. (Particle Data Group), *Review of Particle Physics*, J. Phys. G **37**, 075021 (2010) .

[4] The ATLAS Collaboration, *Observation of $W \to \tau\nu$ Decays with the ATLAS Experiment*, ATLAS-CONF-2010-097, (2010). [http://cdsweb.cern.ch/record/1307529].

[5] The ATLAS Collaboration, *Observation of $Z \to \tau_h\tau_l$ Decays with the ATLAS detector*, ATLAS-CONF-2011-010, (2012). [http://cdsweb.cern.ch/record/1331795].

[6] The 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** (2011) 276–294. 20 p, arXiv:1108.4101 [hep-ex].

[7] The ATLAS Collaboration, *Measurement of the $Z \to \tau\tau$ Cross Section with the ATLAS Detector*, Phys. Rev. D **84** (2011) 112006. 28 p, arXiv:1108.2016 [hep-ex].

[8] The ATLAS Collaboration, *Measurement of Tau Polarization in $W(\to \tau\nu)$+jets Decays with the ATLAS Detector in pp Collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phy. J. C **72** (2012) 1–21, arXiv:1204.6720 [hep-ex].

[9] The ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \to \tau^+\tau^-$ decay mode with $4.7\,fb^{-1}$ of ATLAS data at 7 TeV*, ATLAS-CONF-2012-014, (2012). [http://cdsweb.cern.ch/record/1429662].

[10] The ATLAS Collaboration, *Search for neutral MSSM Higgs bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, ATLAS-CONF-2011-132, (2011). [http://cdsweb.cern.ch/record/138383].

[11] The ATLAS Collaboration, *Search for charged Higgs bosons decaying via $H^+ \to \tau\nu$ in top quark pair events using pp collision data at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **1206** (2012) 039, arXiv:1204.2760 [hep-ex].

[12] The ATLAS Collaboration, *Search for supersymmetry with jets, missing transverse momentum and at least one hadronically decaying tau lepton in proton-proton collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Phys. Lett. B **714** (2012) 197–214, arXiv:1204.3852 [hep-ex].

[13] The ATLAS Collaboration, *Search for Events with Large Missing Transverse Momentum, Jets, and at Least Two Tau Leptons in 7 TeV Proton-Proton Collision Data with the ATLAS Detector*, Phys. Lett. B **714** (2012) 180–196, arXiv:1203.6580 [hep-ex].

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

[15] The ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*, ATLAS-CONF-2011-152, (2011). [http://cdsweb.cern.ch/record/1398195].

[16] 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.

[17] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[18] H. W. Bertini, *Intranuclear-cascade calculation of the secondary nucleon spectra from nucleon-nucleus interactions in the energy range 340 to 2900 mev and comparisons with experiment*, Phys. Rev. **188** (1969) 1711–1730.

[19] S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instr. and Meth. A **506** (2003) 250.

[20] M. Cacciari et al., *The anti-kt jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[21] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, (2008). [http://cdsweb.cern.ch/record/1099735].

[22] T. Barillari et al., *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001-2, 2008. [http://cdsweb.cern.ch/record/1112035].

[23] The ATLAS Collaboration, *Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector*, Eur. Phy. J. C **71** (2011) 1795.

[24] The ATLAS Collaboration, *Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC in 2011*, ATLAS-CONF-2012-054, (2012). [https://cdsweb.cern.ch/record/1453781].

[25] The ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042, (2012). [http://cdsweb.cern.ch/record/1435196].

[26] The ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATL-PHYS-PUB-2011-006, (2011). [http://cdsweb.cern.ch/record/1345327].

[27] The ATLAS Collaboration, *Measurement of the Mis-identification Probability of τ Leptons from QCD Jets and from Electrons*, ATLAS-CONF-2011-113, (2011). [http://cdsweb.cern.ch/record/1375550].

[28] M. Cacciari and G. P. Salam, *Dispelling the N3 myth for the kt jet-finder*, Phys. Lett. **B641** (2006) 57–61, arXiv:hep-ph/0512210.

[29] The ATLAS Collaboration, *Measurement of hadronic tau decay identification efficiency using $W \to \tau\nu$ events*, ATLAS-CONF-2011-093, (2011). [http://cdsweb.cern.ch/record/1365728].

[30] G. Folger and J. P. Wellisch, *String parton models in Geant4*, arXiv:nucl-th/0306007.

[31] B. Andersson, G. Gustafson, and B. Nilsson-Almqvist, *A model for low-pT hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions*, Nuclear Physics B **281** (1987) no. 1-2, 289 – 309.

[32] P. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, arXiv:1005.3457v3 [hep-ph].

[33] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance*, JINST **3** (2008) P02013, http://iopscience.iop.org/1748-0221/3/02/P02013/.

A Tau Reconstruction and Identification Variables

This appendix defines all reconstruction and identification variables used by the jet and electron and muon discriminants. The variables are:

Hadronic radius (R_{Had}): Transverse energy weighted shower width in the hadronic (Had) calorimeter

$$R_{\text{Had}} = \frac{\sum_{i\in\{\text{Had}\}}^{\Delta R_i<0.4} E_{\text{T},i}^{\text{EM}}\,\Delta R_i}{\sum_{i\in\{\text{Had}\}}^{\Delta R_i<0.4} E_{\text{T},i}^{\text{EM}}},$$

where i runs over cells in the hadronic calorimeter and also layer 3 of the EM calorimeter. The energy is calibrated at EM scale. Only cells in a cone of $\Delta R < 0.4$ around the cluster barycentre are considered.

Calorimetic radius (R_{Cal}): Shower width in the electromagnetic (EM) and hadronic (Had) calorimeter weighted by the transverse energy of each calorimeter part.

$$R_{\text{Cal}} = \frac{\sum_{i\in\{\text{EM 0–2}\}}^{\Delta R_i<0.4} E_{\text{T},i}^{\text{EM}}\,\Delta R_i + \sum_{j\in\{\text{Had}\}}^{\Delta R_j<0.4} E_{\text{T},j}^{\text{EM}}\,\Delta R_j}{\sum_{i\in\{\text{EM 0–2}\}}^{\Delta R_i<0.4} E_{\text{T},i}^{\text{EM}} + \sum_{j\in\{\text{Had}\}}^{\Delta R_j<0.4} E_{\text{T},j}^{\text{EM}}},$$

where i runs over cells in the first three layers of the EM calorimeter (pre-sampler, layer 1, and layer 2) and j runs over cells in the hadronic calorimeter and layer 3 of the EM calorimeter. Only cells in a cone of $\Delta R < 0.4$ around the cluster barycentre are considered. All cell energies are calibrated at EM scale.

Track radius (R_{track}): p_{T} weighted track width:

$$R_{\text{track}} = \frac{\sum_{i}^{\Delta R_i<0.4} p_{\text{T},i}\,\Delta R_i}{\sum_{i}^{\Delta R_i<0.4} p_{\text{T},i}},$$

where i runs over all core and isolation tracks of the tau candidate, within $\Delta R_i < 0.4$. ΔR_i is defined relative to the tau jet seed axis and $p_{\text{T},i}$ is the track transverse momentum. Note that for candidates with only one track total in the core cone and isolation annulus, R_{track} simplifies to the ΔR between the track and the axis of the seed jet.

Leading track momentum fraction (f_{track}):

$$f_{\text{track}} = \frac{p_{\text{T},1}^{\text{track}}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.4} E_{\text{T},j}^{\text{EM}}},$$

where $p_{\text{T},1}^{\text{track}}$ is the transverse momentum of the leading p_{T} core track of the tau candidate and j runs over all cells in $\Delta R < 0.4$. The cell calibration at the EM scale is used.

Note that for candidates with one track, f_{track} is the fraction of the candidate's momentum attributed to the track, compared to the total momentum of the candidate, which can have contributions from the calorimeter deposits from π^0s and other neutrals.

Core energy fraction (f_{core}): Fraction of transverse energy in the core ($\Delta R < 0.1$) of the tau candidate:

$$f_{\text{core}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.1} E_{\text{T},i}^{\text{EM}}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.4} E_{\text{T},j}^{\text{EM}}},$$

where i runs over all cells associated to the tau candidate within $\Delta R < 0.1$ of the tau jet seed axis and j runs over all cells in $\Delta R < 0.4$. The cell calibration at the EM scale is used.

Electromagnetic fraction (f_{EM}): Fraction of transverse energy of the tau candidate deposited in the EM calorimeter:

$$f_{\text{EM}} = \frac{\sum_{i\in\{\text{EM 0–2}\}}^{\Delta R_i<0.4} E_{\text{T},i}^{\text{EM}}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.4} E_{\text{T},j}^{\text{EM}}},$$

where $E_{\text{T},i}$ ($E_{\text{T},j}$) is the transverse energy, calibrated at the EM energy scale, deposited in cell i (j), and i runs over the cells in the first three layers of the EM calorimeter, while j runs over the cells in all layers of the electromagnetic calorimeter.

Cluster mass (m_{clusters}): Invariant mass computed from the constituent clusters of the seed jet, calibrated at the LC energy scale.

Track mass (m_{tracks}): Invariant mass of the track system, where the tracks used for the invariant mass calculation use both core and isolation tracks.

Transverse flight path significance ($S_{\text{T}}^{\text{flight}}$): The decay length significance of the secondary vertex for multi-track tau candidates in the transverse plane:

$$S_{\text{T}}^{\text{flight}} = \frac{L_{\text{T}}^{\text{flight}}}{\sigma_{L_{\text{T}}^{\text{flight}}}},$$

where $L_{\text{T}}^{\text{flight}}$ is the reconstructed signed decay length, and $\sigma_{L_{\text{T}}^{\text{flight}}}$ is its estimated uncertainty. Only core tracks are used for the secondary vertex fit.

TRT HT fraction (f_{HT}): The ratio of high-threshold to low-threshold hits (including outlier hits), in the Transition Radiation Tracker (TRT), for the leading p_{T} core track. Since electrons are lighter than pions, and therefore have higher Lorentz γ factors, they are more likely to produce the transition radiation that causes high threshold hits in the TRT [33]. This variable can be used to discriminate hadronic 1-prong tau candidates from electrons.

Ring isolation (f_{iso}):

$$f_{\text{iso}} = \frac{\sum_{i\in\{\text{all}\}}^{0.1<\Delta R<0.2} E_{\text{T},i}^{\text{EM}}}{\sum_{j\in\{\text{all}\}}^{\Delta R<0.4} E_{\text{T},j}^{\text{EM}}},$$

where i runs over calorimeter cells in the associated topocluster of the tau candidate in a annulus within $0.1 < \Delta R < 0.2$ around the seed and j runs over cells in a cone of $\Delta R < 0.4$. The energy is calibrated at electromagnetic scale.

Leading track IP significance ($S_{\text{lead track}}$): impact parameter significance of the leading track of the tau candidate:

$$S_{\text{lead track}} = \frac{d_0}{\sigma_{d_0}},$$

where d_0 is the distance of closest approach of the track to the reconstructed primary vertex in the transverse plane, and σ_{d_0} is its estimated uncertainty.

Maximum ΔR (ΔR_{max}): The maximal ΔR between a track associated to the tau candidate and the tau axis. Only core tracks are considered.

First 2(3) leading cluster energy ratio ($f_{\text{2 lead clusters}}$($f_{\text{3 lead clusters}}$)): the ratio of the energy of the first two (three) leading clusters (highest energy first) over the total energy of all clusters associated to the tau candidate.

Number of tracks in isolation annulus ($N_{\text{trk}}^{0.2<\Delta R<0.4}$): Number of tracks reconstructed within $0.2 < \Delta R < 0.4$ around the seed axis.

Electromagnetic energy over momentum of track system (f_{P}^{EM}):

$$f_{\text{P}}^{\text{EM}} = \frac{\sum_{l=1}^{N_{\text{clus}}} E_l^{\text{EM}}}{\sum_i^{\Delta R<0.2} p_i^{\text{trk}}},$$

where l runs over calorimeter clusters associated to tau candidate, E_l^{EM} denotes the part of cluster energy deposited in the electromagnetic part of each cluster (presampler and first two layers) and i runs over tracks associated to tau candidate in the core region. The cluster energy is calibrated at the LC scale.

Presampler strip energy fraction (f_{PS}):

$$f_{\text{PS}} = \frac{\sum_{l=1}^{N_{\text{clus}}} E_l^{\text{PS}}}{\sum_{l=1}^{N_{\text{clus}}} E_l},$$

where l runs over calorimeter clusters associated to tau candidate, E_l^{PS} denotes the part of cluster energy, calibrated at the LC scale, deposited in the Presampler layer of calorimeter and E_l is the total energy of a calorimeter cluster.

Electromagnetic energy of charged pions over calorimetric electromagnetic energy ($f_{\text{EM}}^{\pi^{\pm}}$):

$$f_{\text{EM}}^{\pi^{\pm}} = \frac{\sum_i^{\Delta R<0.2} p_i^{trk} - \sum_{l=1}^{N_{\text{clus}}} E_l^{\text{Had}}}{\sum_{l=1}^{N_{\text{clus}}} E_l^{\text{EM}}},$$

where l runs over calorimeter clusters associated to tau candidate, E_l^{Had} denotes the part of cluster energy deposited in the hadronic part of each cluster, E_l^{EM} is the part of cluster energy deposited in the electromagnetic part of each cluster (presampler and first two layers) and i runs over tracks associated to tau candidate in the core region. All clusters are calibrated at the LC energy scale.