Performance of reconstruction and identification of $\tau$ leptons decaying to hadrons and $\nu_\tau$ in pp collisions at $\sqrt{s}=13$ TeV

To cite this article: A.M. Sirunyan et al 2018 JINST 13 P10005

View the article online for updates and enhancements.
Performance of reconstruction and identification of $\tau$ leptons decaying to hadrons and $\nu_\tau$ in $pp$ collisions at $\sqrt{s} = 13$ TeV

The CMS collaboration

E-mail: cms-publication-committee-chair@cern.ch

ABSTRACT: The algorithm developed by the CMS Collaboration to reconstruct and identify $\tau$ leptons produced in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV, via their decays to hadrons and a neutrino, has been significantly improved. The changes include a revised reconstruction of $\pi^0$ candidates, and improvements in multivariate discriminants to separate $\tau$ leptons from jets and electrons. The algorithm is extended to reconstruct $\tau$ leptons in highly Lorentz-boosted pair production, and in the high-level trigger. The performance of the algorithm is studied using proton-proton collisions recorded during 2016 at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of $35.9\,\text{fb}^{-1}$. The performance is evaluated in terms of the efficiency for a genuine $\tau$ lepton to pass the identification criteria and of the probabilities for jets, electrons, and muons to be misidentified as $\tau$ leptons. The results are found to be very close to those expected from Monte Carlo simulation.

KEYWORDS: Large detector-systems performance; Particle identification methods; Performance of High Energy Physics Detectors
## Contents

1 Introduction 1

2 The CMS detector 2

3 Data and simulated events 3

4 Event reconstruction 4

5 Reconstruction and identification of $\tau_h$ 5
   5.1 The hadrons-plus-strips algorithm 5
      5.1.1 Dynamic strip reconstruction 6
      5.1.2 Mass constraints on decay modes 8
   5.2 Discrimination of $\tau_h$ candidates against jets 9
      5.2.1 Isolation sum discriminants 9
      5.2.2 MVA-based discriminants 10
   5.3 Discrimination of $\tau$ leptons against electrons 13
   5.4 Discrimination of $\tau$ leptons against muons 14

6 Reconstruction of highly boosted $\tau$ lepton pairs 15

7 Identification of $\tau_h$ candidates in the high-level trigger 16

8 Event selection and systematic uncertainties 19
   8.1 The $Z/\gamma^* \rightarrow \tau\tau$ events 19
   8.2 The $\mu\tau_h$ final states in $t\bar{t}$ events 19
   8.3 The $Z/\gamma^* \rightarrow \mu\mu$ events to constrain the $Z/\gamma^* \rightarrow \ell\ell$ normalization 20
   8.4 Off-shell $W \rightarrow \tau\nu$ events 20
   8.5 Off-shell $W \rightarrow \mu\nu$ events to constrain the $W \rightarrow \tau\nu$ normalization 20
   8.6 Events from $W \rightarrow \mu\nu+$jet production 20
   8.7 The $e\mu$ final states in $t\bar{t}$ events 21
   8.8 The $Z/\gamma^* \rightarrow ee, \mu\mu$ events for measuring the $e/\mu \rightarrow \tau_h$ misidentification probability 21
   8.9 Systematic uncertainties affecting all studied final states 21

9 Measurement of the $\tau_h$ identification efficiency 22
   9.1 Using the tag-and-probe method in $Z/\gamma^*$ events 23
   9.2 Using $t\bar{t}$ events 25
   9.3 Using off-shell $W \rightarrow \tau\nu$ events 27
   9.4 Extrapolation of the $\tau_h$ identification efficiency to large $p_T^{\tau_h}$ 29
   9.5 Using the tag-and-probe method in $Z/\gamma^*$ events for highly boosted $\tau$ lepton pairs 29
1 Introduction

Searches for new phenomena that consider signatures with \( \tau \) leptons have gained great interest in proton-proton (pp) collisions at the CERN LHC. The most prominent one among these is the decay of Higgs bosons (H) to pairs of \( \tau \) leptons, which constitutes an especially sensitive channel for probing Higgs boson couplings to fermions. The observation of the standard model (SM) Higgs boson decaying to a pair of \( \tau \) leptons has recently been reported [1, 2]. Moreover, searches with \( \tau \) leptons in the final state have high sensitivity to the production of both neutral and charged Higgs bosons expected in the minimal supersymmetric standard model (MSSM) [3, 4], in which enhancements in the couplings to \( \tau \) leptons can be substantial at large \( \tan \beta \), where \( \tan \beta \) is the ratio of vacuum expectation values of the two Higgs doublets in the MSSM. Examples of such searches can be found in refs. [5–7]. In addition, searches for particles beyond the SM, such as new or heavy Higgs bosons [8–11], leptoquarks [12], supersymmetric particles [13–16], or gauge bosons [17–19] benefit significantly from any improvements made in \( \tau \) lepton reconstruction and identification.

The \( \tau \) lepton, with a mass of \( m_\tau = 1776.86 \pm 0.12 \text{ MeV} \) [20], is the only lepton sufficiently massive to decay into hadrons and a neutrino. About one third of the time, \( \tau \) leptons decay into an electron or a muon, and two neutrinos. The neutrinos escape undetected, but the e and \( \mu \) are reconstructed and identified through the usual techniques available for such leptons [21, 22]. These decay final states are denoted as \( \tau_e \) and \( \tau_\mu \), respectively. Almost all the remaining decay final states of \( \tau \) leptons contain hadrons, typically with a combination of charged and neutral mesons, and a \( \nu_\tau \).

The decays of \( \tau \) leptons into hadrons and neutrinos, denoted by \( \tau_h \), are reconstructed and identified using the hadrons-plus-strips (HPS) algorithm [23, 24], which was developed and used in CMS when the LHC operated at \( \sqrt{s} = 7 \) and 8 TeV. The HPS algorithm reconstructs the \( \tau_h \) modes by combining information from charged hadrons, which are reconstructed using their associated tracks in the inner tracker, and \( \pi^0 \) candidates, obtained by clustering photon and electron candidates from photon conversions in rectangular regions of pseudorapidity and azimuth, \( \eta \times \phi \) regions, called

\begin{itemize}
\item 10 Jet \( \rightarrow \tau_h \) misidentification probability
\item 10.1 Using \( W \rightarrow \mu \nu + \text{jet} \) events
\item 10.2 Using \( e\mu + \text{jets} \) events
\item 11 The \( e/\mu \rightarrow \tau_h \) misidentification probability
\item 11.1 Measurement of the \( e \rightarrow \tau_h \) probability
\item 11.2 Measurement of the \( \mu \rightarrow \tau_h \) probability
\item 12 Measurement of the \( \tau_h \) energy scale
\item 13 Performance of \( \tau_h \) identification in the high-level trigger
\item 14 Summary
\end{itemize}

The CMS collaboration

\[ 1 \]
“strips”. The major challenge in the identification of $\tau_h$ is to distinguish these objects from quark and gluon jets, which are copiously produced in pp collisions. The primary method for reducing backgrounds from jets misidentified as $\tau_h$ candidates exploits the fact that there are fewer particles present in $\tau_h$ decays, and that their energies are deposited in narrow regions of ($\eta, \phi$) compared to those from energetic quark or gluon jets. In certain analyses, the misidentification (MisID) of electrons or muons as $\tau_h$ candidates can also constitute a sizeable background.

The $\tau_h$ identification algorithm improved for analyzing data at $\sqrt{s} = 13$ TeV contains the following new features:

1. A modification of the strip reconstruction algorithm, to the so-called dynamic strip reconstruction, that changes the size of a strip in a dynamic way that collects the $\pi^0$ decay products more effectively;
2. improvements in the multivariate-analysis (MVA) based discriminant [24] that reduces the background from jets, by combining information on isolation, lifetime of the $\tau$ lepton, and energy distribution in the shower; and
3. improvements in the MVA-based discriminant that suppresses electrons misidentified as $\tau_h$ candidates.

This paper is organized as follows. After a brief introduction of the CMS detector in section 2, we discuss the data and the event simulations used to evaluate the performance of the HPS algorithm in section 3. The reconstruction and identification of physical objects (other than $\tau_h$) is briefly described in section 4. Section 5 describes the HPS algorithm used for 13 TeV data and its simulation. The extended version of the algorithm used to reconstruct $\tau_h$ pairs produced in topologies with high Lorentz-boosts is presented in section 6, while the specialized version developed for trigger purposes is discussed in section 7. The selection of events used to evaluate the performance of the $\tau_h$ reconstruction algorithm, as well as systematic uncertainties common to all measurements are discussed in section 8. The performance evaluation of the improved algorithm using selected data samples is given thereafter: section 9 describes the $\tau_h$ identification efficiency, while sections 10 and 11 summarize the respective jet $\rightarrow \tau_h$ and $e/\mu \rightarrow \tau_h$ misidentification probabilities. The $\tau_h$ energy scale is discussed in section 12. Finally, section 13 presents the performance of $\tau_h$ identification in the high-level trigger, and a brief summary in section 14 concludes this paper.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8T. A silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections, reside within the field of the solenoid. Extensive forward calorimetry complements the coverage provided by the barrel and endcaps. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid.

The CMS tracker is a cylindrical detector, constructed from 1440 silicon-pixel and 15 148 silicon-strip detector modules that cover the range of $|\eta| < 2.5$. Tracks of charged hadrons are reconstructed with typical efficiencies of 80–90%, depending on transverse momentum ($p_T$) and
The silicon tracker presents a significant amount of material in front of the ECAL, mostly due to the mechanical structure, the associated services, and the cooling system. A minimum of 0.4 radiation lengths ($X_0$) of material is present at $|\eta| \approx 0$, which rises to $\approx 2.0 X_0$ at $|\eta| \approx 1.4$, and decreases to $\approx 1.3 X_0$ at $|\eta| \approx 2.5$. Photons originating from $\pi^0$ decays therefore have a high probability to convert into $e^-e^+$ pairs within the volume of the tracker.

The ECAL is a homogeneous and hermetic calorimeter made of PbWO$_4$ scintillating crystals. It is composed of a central barrel, covering the region $|\eta| < 1.48$, and two endcaps, covering $1.48 < |\eta| < 3.0$. The small radiation length ($X_0 = 0.89$ cm) and small Molière radius (2.3 cm) of the PbWO$_4$ crystals provide a compact calorimeter with excellent two-shower separation. The ECAL is $> 25 X_0$ thick.

The HCAL is a sampling calorimeter made of brass and plastic scintillator, with a coverage up to $|\eta| = 3.0$. The scintillation light is converted by wavelength-shifting fibres and channelled to photodetectors via clear fibres. The thickness of the HCAL is in the range 7–11 interaction lengths, depending on $\eta$.

The muon detection system is made up of four planes of gas-ionization detectors, where each plane consists of several layers of aluminium drift tubes (DTs) in the barrel region and cathode strip chambers (CSCs) in the endcap region, complemented by resistive-plate chambers (RPCs) that are used only in the trigger.

A two-tiered trigger system [27] is employed to select interesting events from the LHC bunch crossing rate of up to 40 MHz. The first level (L1), composed of custom-made hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of $\approx 100$ kHz, within a fixed time interval of less than 4 $\mu$s. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software, optimized for fast processing, and reduces the event rate to $\approx 1$ kHz before data storage.

A more detailed description of the CMS detector, together a definition of the coordinate system and kinematic variables, can be found in ref. [28].

3 Data and simulated events

The performance of $\tau_h$ reconstruction and identification algorithms are evaluated in pp collisions recorded by CMS during 2016 at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 35.9 fb$^{-1}$. The Monte Carlo (MC) simulated signal samples contain $H \rightarrow \tau\tau$, $Z' \rightarrow \ell\ell$, $W' \rightarrow \ell \nu$, and $Z/\gamma^* \rightarrow \ell\ell$ events, where $\ell$ refers to $e$, $\mu$, or $\tau$ leptons. Simulated signal contributions from $H \rightarrow \tau\tau$, $Z' \rightarrow \ell\ell$ (with masses up to 4 TeV), $W' \rightarrow \ell \nu$ (with masses up to 5.8 TeV), and MSSM $H \rightarrow \tau\tau$ (with masses up to 3.2 TeV) are used to optimize the identification of $\tau_h$ candidates over a wide range of their $p_T$ values. The $H \rightarrow \tau\tau$ events are generated at next-to-leading order (NLO) in perturbative quantum chromodynamics (QCD) using POWHEG v2 [29–33], while $Z'$ and $W'$ boson events are generated using leading-order (LO) PYTHIA 8.212 [34]. In simulation, the reconstructed $\tau_h$ candidate is taken as matched to the generated $\tau_h$ when both objects lie within a cone of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.3$, where $\Delta \phi$ and $\Delta \eta$ are the distances respectively in $\phi$ and $\eta$ between the reconstructed and generated candidates.

The $W$+jets and $Z/\gamma^* \rightarrow \ell\ell$ events are generated at LO in perturbative QCD using MADGRAPH5_aMC@NLO v2.2.2 [35] with the MLM jet merging scheme [36], while the single top quark
and $t \bar{t}$ events are generated at NLO in perturbative QCD using \textsc{powheg} \cite{37–39}. The diboson \(WW\), \(WZ\), and \(ZZ\) events are generated at NLO using \textsc{MadGraph}5 \textsc{aMC@nlo} with the FXFX jet merging scheme \cite{40} or \textsc{powheg} \cite{41}, while events comprised uniquely of jets produced through the strong interaction, referred to as QCD multijet events, are generated at LO with \textsc{pythia}. The \textsc{pythia} generator, with the CUETP8M1 underlying-event tune \cite{42}, is used to model the parton shower and hadronization processes, as well as $\tau$ lepton decays in all events. The $Z/\gamma^* \rightarrow \ell\ell$ and $W +$jets samples are normalized according to cross sections computed at next-to-next-to-leading order (NNLO) in perturbative QCD accuracy \cite{43–47}, while the $t \bar{t}$ sample is normalised to the cross section computed at NNLO supplemented by soft-gluon resummation with next-to-next-to-leading logarithmic accuracy \cite{48, 49}. The cross sections for single top quark and diboson production are computed at NLO in perturbative QCD accuracy \cite{50}. The production of off-shell W bosons ($m_W > 200$ GeV), with subsequent $W \rightarrow \tau\nu$ or $W \rightarrow \mu\nu$ decays, is simulated at LO with the \textsc{pythia} generator. The differential cross section is reweighted as a function of the invariant mass of the W boson decay products, incorporating NNLO QCD and NLO electroweak corrections \cite{46, 51, 52}. The NNPDF3.0 parton distribution functions \cite{53} are used in all the calculations.

Additional pp collisions that overlap temporally the interactions of interest, referred to as pileup (PU), are generated using \textsc{pythia}, and overlaid on all MC events according to the luminosity profile of the analyzed data. The generated events are passed through a detailed simulation of the CMS detector based on \textsc{Geant4} \cite{54}, and are reconstructed using the same CMS reconstruction software as used for data.

4 Event reconstruction

The particles emerging from pp collisions, such as charged and neutral hadrons, photons, electrons, and muons, are reconstructed and identified by combining the information from the CMS subdetectors using a particle-flow (PF) algorithm \cite{55}. These particles are further grouped to reconstruct higher-level objects, such as jets, missing transverse momentum, $\tau_h$ candidates, and to quantify lepton isolation.

The trajectories of charged particles are reconstructed from their hits in the silicon tracker \cite{26}, and are referred to as tracks.

Electrons are reconstructed from their trajectories in the tracker and from clusters of energy deposition in the ECAL \cite{21}. Electron identification relies on the energy distribution in the electromagnetic shower and on other observables based on tracker and calorimeter information. The selection criteria depend on the $p_T$ and $|\eta|$ of the electron, and on a categorization according to observables sensitive to the amount of bremsstrahlung emitted along the trajectory in the tracker.

Muons are reconstructed by combining tracks reconstructed in both the inner tracker and the outer muon spectrometer \cite{22}. The identification of muons is based on the quality criteria of reconstructed muon tracks, and through requirements of minimal energy deposition along the muon track in the calorimeters.

The isolation of individual electrons or muons ($I_{el/\mu}$) is measured relative to their transverse momenta $p_T^{el/\mu}$ by summing over the scalar $p_T$ values of charged and neutral hadrons, as well as photons, in a cone of $\Delta R < 0.3$ for electrons or 0.4 for muons around the direction of the lepton at
the interaction vertex:

$$I_{e/\mu} = \left( \sum p_T^{\text{charged}} + \max \left[ 0, \sum p_T^{\text{neutral}} + \sum p_T^{\gamma} - p_T^{\text{PU}} \right] \right) / p_T^{e/\mu}. \quad (4.1)$$

The primary pp interaction vertex is defined as the reconstructed vertex with largest value of summed $p_T^2$ of jets, clustered using all tracks assigned to the vertex, and of the associated missing transverse momentum, taken as the negative vector sum of the $p_T$ of those jets. To suppress the contribution from PU, the charged hadrons are required to originate from the primary vertex. The neutral contribution to the isolation from PU (referred to as $p_T^{\text{PU}}$) is estimated through a jet area method [56] for electrons. For muons, the $p_T^{\text{PU}}$ contribution is estimated using the sum of the scalar $p_T$ of charged hadrons not originating from the primary vertex, scaled down by a factor of 0.5 (to accommodate the assumed ratio for the production of neutral and charged hadrons).

Jets are clustered from PF particles using the infrared and collinear-safe anti-$k_T$ algorithm [57, 58] with a distance parameter of 0.4. The jet momentum is defined by the vectorial sum of all particle momenta in the jet. The simulation is found to provide results for jet $p_T$ within 5 to 10% of their true values over the whole $p_T$ spectrum and detector acceptance. To suppress contributions from PU, charged hadrons not originating from the primary vertex are discarded, and an offset correction is applied to correct the remaining PF contributions. Jet energy corrections are obtained from simulation to bring the measured response of jets to that of particle level jets on average, and are confirmed with in situ measurements through momentum balance in dijet, $\gamma$+jet, $Z$+jet, and multijet events [59]. The combined secondary vertex v2 (CSVv2) b tagging algorithm [60] with a medium working point (WP) is used to identify jets originating from b quarks. The working point corresponds to an identification efficiency of about 70% for b quark jets with $p_T > 30$ GeV, and a probability for light-quark or gluon jets to be misidentified as b quarks of $\approx 1\%$.

The missing transverse momentum vector, $\vec{p}_T^{\text{miss}}$, is defined as the projection of the negative vector sum of the momenta of all reconstructed particles in an event on the plane perpendicular to the beams. The $\vec{p}_T^{\text{miss}}$ is corrected by propagating to it all the corrections made to the momenta of jets. Its magnitude is referred to as $p_T^{\text{miss}}$.

5 Reconstruction and identification of $\tau_h$

The basic features of the HPS algorithm are identical to those used during the previous data taking at $\sqrt{s} = 7$ and 8 TeV [24], except for the improvements in $\pi^0$ reconstruction described below in section 5.1.1. Sections 5.2, 5.3, and 5.4 discuss the discriminants used to distinguish reconstructed $\tau_h$ candidates from jets, electrons, and muons, respectively.

5.1 The hadrons-plus-strips algorithm

Starting from the constituents of reconstructed jets, the HPS algorithm reconstructs the different decays of the $\tau$ lepton into hadrons. The final states include charged hadrons, as well as neutral pions, as shown in table 1. The $\pi^0$ mesons promptly decay into pairs of photons, which have a high probability of converting into $e^+e^-$ pairs as they traverse the tracker material. The large magnetic field of the CMS solenoid leads to a spatial separation of the $e^+e^-$ pairs in the $(\phi, \eta)$ plane. To reconstruct the full energy of the neutral pions, the electron and photon candidates falling within
a certain region of $\Delta \eta \times \Delta \phi$ are clustered together, with the resulting object referred to as a “strip”. The strip momentum is defined by the vectorial sum of all its constituent momenta. The procedure is described in section 5.1.1, together with the improvements introduced to the previous algorithm.

Table 1. Weak decays of $\tau$ leptons and their branching fractions ($B$) in % [20] are given, rounded to one decimal place. Also, where appropriate, we indicate the known intermediate resonances of all the listed hadrons. Charged hadrons are denoted by the symbol $h^\pm$. Although for simplicity we show just $\tau^-$ decays in the table, the values are also valid for the charge-conjugate processes.

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>Resonance</th>
<th>$B$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leptonic decays</td>
<td>35.2</td>
<td></td>
</tr>
<tr>
<td>$\tau^- \rightarrow e^- \bar{\nu}<em>e \nu</em>\tau$</td>
<td>17.8</td>
<td></td>
</tr>
<tr>
<td>$\tau^- \rightarrow \mu^- \bar{\nu}<em>\mu \nu</em>\tau$</td>
<td>17.4</td>
<td></td>
</tr>
<tr>
<td>Hadronic decays</td>
<td>64.8</td>
<td></td>
</tr>
<tr>
<td>$\tau^- \rightarrow h^- \nu_\tau$</td>
<td>11.5</td>
<td></td>
</tr>
<tr>
<td>$\tau^- \rightarrow h^- \pi^0 v_\tau$</td>
<td>$\rho(770)$</td>
<td>25.9</td>
</tr>
<tr>
<td>$\tau^- \rightarrow h^- \pi^0 \pi^0 v_\tau$</td>
<td>$a_1(1260)$</td>
<td>9.5</td>
</tr>
<tr>
<td>$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$</td>
<td>$a_1(1260)$</td>
<td>9.8</td>
</tr>
<tr>
<td>$\tau^- \rightarrow h^- h^+ h^0 \nu_\tau$</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3.3</td>
<td></td>
</tr>
</tbody>
</table>

Charged particles used in the reconstruction of $\tau_h$ candidates are required to have $p_T > 0.5$ GeV, and must be compatible with originating from the primary vertex of the event, where the criterion on the transverse impact parameter is not highly restrictive ($d_{xy} < 0.1$ cm), to minimize the rejection of genuine $\tau$ leptons with long lifetimes. The requirement of $p_T > 0.5$ GeV on the charged particles ensures that the corresponding tracks have sufficient quality, and pass a minimal requirement on the number of layers with hits in the tracking detector.

Based on the set of charged particles and strips contained in a jet, the HPS algorithm generates all possible combinations of hadrons for the following decay modes: $h^\pm$, $h^+\pi^0$, $h^+\pi^0\pi^0$, and $h^+h^+h^-$. The reconstructed mass of the “visible” hadronic constituents of the $\tau_h$ candidate (i.e., the decay products, excluding neutrinos) is required to be compatible either with the $\rho(770)$, or the $a_1(1260)$ resonances in the $h^+\pi^0$ and in the $h^+\pi^0\pi^0$ or $h^+h^+h^-$ decay modes, respectively, as discussed in section 5.1.2. The $h^+\pi^0$ and $h^+\pi^0\pi^0$ modes are consolidated into the $h^+\pi^0$ mode, and are analyzed together. The combinations of charged particles and strips considered by the HPS algorithm represent all the hadronic $\tau$ lepton decay modes in table 1, except $\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$ with $B = 4.8\%$. This decay is not considered in the current version of the algorithm, because of its greater contamination by jets. The $\tau_h$ candidates of charge other than $\pm 1$ are rejected, as are those with charged particles or strips outside the signal cone, defined by $R_{\text{sig}} = (3.0\text{GeV})/p_T$, where the $p_T$ is that of the hadronic system, with cone size limited to the range 0.05–0.10. Finally, only the $\tau_h$ candidate with largest $p_T$ is kept for further analysis, resulting in a single $\tau_h$ candidate per jet.

### 5.1.1 Dynamic strip reconstruction

Photon and electron constituents of jets, which seed the $\tau_h$ reconstruction, are clustered into $\Delta \eta \times \Delta \phi$ strips, and used to collect all energy depositions in the ECAL that arise from neutral pions produced.
in $\tau_h$ decays. The size of the $\Delta\eta \times \Delta\phi$ window is set to a fixed value of $0.05 \times 0.20$ in the ($\eta$, $\phi$) plane in the previous version of the HPS algorithm [24]. However, this fixed strip size is not always adequate to contain all electrons and photons originating from the $\tau_h$ decays, meaning that some of the particles from $\tau_h$ lepton decay can contribute to the isolation region and thereby reduce the isolation efficiency for genuine $\tau_h$ candidates.

Our studies of $\tau_h$ reconstruction have led to the following observations:

1. A charged pion from $\tau_h$ decays undergoing nuclear interactions in the tracker material can produce secondary particles with lower $p_T$. This can result in cascades of low-$p_T$ electrons and photons that can appear outside of the strip window, and affect the isolation of a $\tau_h$ candidate, despite these particles originating from remnants of the $\tau_h$ decay.

2. Photons from $\pi^0$ decays have a large probability to convert into $e^+e^-$ pairs and, after multiple scattering and bremsstrahlung, some of the remaining electrons and photons can end up outside a fixed size window, also affecting the isolation.

Naively, these decay products can be integrated into the strip by suitably increasing its size. Conversely, if the $\tau_h$ has large $p_T$, the decay products tend to be boosted in the direction of the $\tau_h$ candidate momentum. In this case, a smaller than previously considered strip size can reduce background contributions to that strip, while taking full account of all decay products.

Based on these considerations, the strip clustering of the HPS algorithm has been changed as follows:

1. The electron or photon ($e/\gamma$) with the highest $p_T$ not yet included in any strip is used to seed a new strip, with initial position set to the $\eta$ and $\phi$ values of the new $e/\gamma$ seed.

2. The $p_T$ of the second-highest $e/\gamma$ deposition within

   \[ \Delta \eta = f(p_T^{e/\gamma}) + f(p_T^{\text{strip}}) \quad \text{and} \quad \Delta \phi = g(p_T^{e/\gamma}) + g(p_T^{\text{strip}}) \]

   of the strip position is merged into the strip. The dimensionless functions $f$ and $g$ are determined from single $\tau$ lepton events, generated in MC with uniform $p_T$ in the range from 20 to 400 GeV and $|\eta| < 2.3$, such that 95% of all electrons and photons that arise from $\tau_h$ decays are contained within one strip. The functional form is based on the $\Delta\eta$ and $\Delta\phi$ between the $\tau_h$ and the $e/\gamma$ candidate, studied as a function of the $p_T$ of the $e/\gamma$ candidate. As shown in figure 1, the 95% envelope of points in each bin is fitted using the analytic form $a/(p_T)^b$, resulting in:

   \[ f(p_T) = 0.20 p_T^{-0.66} \quad \text{and} \quad g(p_T) = 0.35 p_T^{-0.71}, \]

   where the $p_T$ is in GeV. The upper limits on the strip size are set to 0.3 in $\Delta\phi$ and 0.15 in $\Delta\eta$, and the lower limits are set to 0.05 for both $\Delta\phi$ and $\Delta\eta$. The size of the window depends on the $p_T$ values of both the strip and the merged $e/\gamma$ candidate. The size is defined by the maximum separation between the two objects, assuming they have opposite charges and are produced back-to-back in their rest frame. Although, strictly speaking, this reasoning applies only to the $\phi$ direction, it is also used for the $\eta$ direction.
3. The strip position is recomputed using the $p_T$-weighted average of all $e/\gamma$ constituents in the strip:

\[
\eta_{\text{strip}} = \frac{1}{P_{\text{strip}}^T} \sum_{e/\gamma} p_{e/\gamma}^T \eta_{e/\gamma}, \\
\phi_{\text{strip}} = \frac{1}{P_{\text{strip}}^T} \sum_{e/\gamma} p_{e/\gamma}^T \phi_{e/\gamma}.
\]  

(5.3)

4. The construction of the strip ends when there is no other $e/\gamma$ candidate within the $\Delta \eta \times \Delta \phi$ window. In this case, the clustering proceeds by selecting a new strip, seeded by the $e/\gamma$ candidate of highest $p_T$ that is not as yet associated with any strip.

As defined above, the size of the strip does not depend on the cone-size of the $\tau_h$ signal. The $p_T$-weighted center ($\eta, \phi$) of the strip is required to be within the signal cone, while part of the strip can lie outside of it.

![Figure 1](image_url)

**Figure 1.** Distance in $\eta$ (left) and in $\phi$ (right) between the $\tau_h$ and $e/\gamma$ candidates for $\tau_h$ decay products, as a function of the $p_T$ of the $e/\gamma$ candidate, in simulated $\tau_h$ decays. The points show the 95% envelope for a given bin, and the solid red lines represent the fitted functions $f$ and $g$ given in eq. (5.2).

### 5.1.2 Mass constraints on decay modes

Strips are combined with charged particles to form $\tau_h$ decay hypotheses. Then, to check the compatibility of each hypothesis with the signatures expected from different $\tau_h$ decay modes, the reconstructed mass of the visible hadronic constituents of the $\tau_h$ candidate (that we refer to as $m_{\tau_h}$) is required to lie within a mass window corresponding either to the $\rho(770)$ or $a_1(1260)$ meson. The widths and positions of the mass windows are optimized for each decay mode to maximize the ratio of the $\tau_h$ reconstruction efficiency to the jet $\rightarrow \tau_h$ misidentification probability, with results that can be summarized as follows:

1. $0.3 \text{ GeV} - \Delta m_{\tau_h} < m_{\tau_h} < 1.3 \text{ GeV} \sqrt{p_T^h/(100 \text{ GeV})} + \Delta m_{\tau_h}$ for $h^+\pi^0$, with the mass window enlarged for $\tau_h$ candidates of high $p_T$ to account for resolution, and the upper limit on the mass window constrained to lie between 1.3 and 4.2 GeV,
2. 0.4 GeV $- \Delta m_{\tau_{\text{h}}} < m_{\tau_{\text{h}}} < 1.2$ GeV $\sqrt{p_{\text{T}}^{\tau_{\text{h}}}/(100\text{GeV})} + \Delta m_{\tau_{\text{h}}}$ for $h^+\pi^0\pi^0$, with the upper limit on the mass window restricted to lie between 1.2 and 4.0 GeV, and

3. $0.8 < m_{\tau_{\text{h}}} < 1.5$ GeV for the $h^+h^0h^0$ channels,

where $\Delta m_{\tau_{\text{h}}}$ is the change in the mass of the $\tau_{\text{h}}$ candidate brought about by the addition of the $e/\gamma$ candidates to its strip. It is calculated as follows:

$$\Delta m_{\tau_{\text{h}}} = \sqrt{\left(\frac{\partial m_{\tau_{\text{h}}}}{\partial \eta_{\text{strip}}} f(p_{T}^{\text{strip}})\right)^2 + \left(\frac{\partial m_{\tau_{\text{h}}}}{\partial \phi_{\text{strip}}} g(p_{T}^{\text{strip}})\right)^2}, \quad (5.4)$$

with

$$\frac{\partial m_{\tau_{\text{h}}}}{\partial \eta_{\text{strip}}} = \frac{p_{z}^{\text{strip}} E_{\tau_{\text{h}}} - E_{\text{strip}} p_{z}^{\text{strip}}}{m_{\tau_{\text{h}}}}, \quad \text{and}$$

$$\frac{\partial m_{\tau_{\text{h}}}}{\partial \phi_{\text{strip}}} = -\left(p_{y}^{\tau_{\text{h}}} - p_{y}^{\text{strip}}\right) p_{x}^{\text{strip}} + \left(p_{z}^{\tau_{\text{h}}} - p_{z}^{\text{strip}}\right) p_{y}^{\text{strip}}.$$

where $p_{\tau_{\text{h}}} = (E_{\tau_{\text{h}}}, p_{x}^{\tau_{\text{h}}, y, z} \tau_{\text{h}} )$ and $p_{\text{strip}} = (E_{\text{strip}}, p_{x}^{\text{strip}, y, z} )$ are the four-momenta of the $\tau_{\text{h}}$ and of the strip, respectively.

### 5.2 Discrimination of $\tau_{\text{h}}$ candidates against jets

Requiring $\tau_{\text{h}}$ candidates to pass certain specific isolation requirements provides a strong handle for reducing the jet $\rightarrow \tau_{\text{h}}$ misidentification probability. The two $\tau_{\text{h}}$ isolation discriminants developed previously [24], namely the isolation sum and the MVA-based discriminants, have now been reoptimized. A cone with $\Delta R = 0.5$ was originally used in the definition of isolation for all event types. However, in processes with a high number of final-state objects, such as for Higgs boson production in association with top quarks ($t\bar{t}H$), the isolation is affected by the presence of nearby objects. Studies using such $t\bar{t}H$ events with $H \rightarrow \tau\tau$ decays led to the conclusion that a smaller isolation cone improves the $\tau_{\text{h}}$ efficiency in such events. A smaller isolation cone of radius $\Delta R = 0.3$ is therefore now used in these types of events.

#### 5.2.1 Isolation sum discriminants

The isolation of $\tau_{\text{h}}$ candidates is computed by summing the scalar $p_{T}$ of charged particles ($\sum p_{T}^{\text{charged}}$) and photons ($\sum p_{T}^{\gamma}$) reconstructed using the PF algorithm within the isolation cone centered on the direction of the $\tau_{\text{h}}$ candidate. Charged-hadron and photon constituents of $\tau_{\text{h}}$ candidates are excluded from the $p_{T}$ sum, defining thereby the isolation as:

$$I_{\tau_{\text{h}}} = \sum p_{T}^{\text{charged}}(d_{z} < 0.2 \text{ cm}) + \max \left(0, \sum p_{T}^{\gamma} - \Delta \beta \sum p_{T}^{\text{charged}}(d_{z} > 0.2 \text{ cm})\right). \quad (5.5)$$

The contribution from PU is suppressed by requiring the charged particles to originate from the production vertex of the $\tau_{\text{h}}$ candidate within a distance of $d_{z} < 0.2$ cm. The PU contribution to the $p_{T}$ sum of photons in the isolation cone is estimated by summing the scalar $p_{T}$ of charged particles not originating from the vertex of the $\tau_{\text{h}}$ candidate ($\sum p_{T}^{\text{charged}}$ with $d_{z} > 0.2$ cm), but appearing within a cone of $\Delta R = 0.8$ around the $\tau_{\text{h}}$ direction multiplied by a so-called $\Delta \beta$ factor, which
accounts for the ratio of energies carried by charged hadrons and photons in inelastic pp collisions, as well as for the different cone sizes used to estimate the PU contributions.

Previously, an empirical factor of 0.46 was used as the $\Delta \beta$ [24]. However, this is found to overestimate the PU contribution to the isolation in data taken in 2015 and 2016. And a new $\Delta \beta$ factor of 0.2 is therefore chosen. This value corresponds approximately to the ratio of neutral to charged pion production rates (0.5), corrected for the difference in the size of the isolation cone ($\Delta R = 0.5$) and the cone used to compute the $\Delta \beta$ correction ($\Delta R = 0.8$): $0.5 \times (0.5^2/0.8^2) \approx 0.195$.

The loose, medium, and tight working points of the isolation sum discriminants are defined by requiring $I_{\eta\tau}$ to be less than 2.5, 1.5, or 0.8 GeV, respectively. These thresholds are chosen such that the resulting efficiencies for the three working points cover the range required for the analyses.

In dynamic strip reconstruction, a photon candidate outside the signal cone can still contribute to the signal. This effectively increases the jet $\rightarrow \tau_h$ misidentification probability because of the decrease in the value of $I_{\eta\tau}$ for misidentified $\tau_h$ candidates. An additional handle is therefore exploited to reduce the jet $\rightarrow \tau_h$ misidentification probability using the scalar $p_T$ sum of $e/\gamma$ candidates included in strips, but located outside of the signal cone, which is defined as

$$p_T^{\text{strip, outer}} = \sum p_T^{e/\gamma} (\Delta R > R_{\text{sig}}).$$

A reduction of about 20% in the jet $\rightarrow \tau_h$ misidentification probability is achieved by requiring $p_T^{\text{strip, outer}}$ to be less than 10% of $p_T^{\tau_h}$, for similar values of efficiency.

A comparison of the expected performance of the isolation sum discriminant for the previous and current versions of the HPS algorithm is shown in figure 2. The efficiency is calculated for generated $\tau_h$ candidates with $p_T > 20$ GeV, $|\eta| < 2.3$, having a decay mode of $h^\pm h^\mp h$, $h^0 h^0 h^0$, or $h^0 h^0 h^0$, and matching to a reconstructed $\tau_h$ candidate with $p_T > 18$ GeV. The misidentification probability is calculated for jets with $p_T > 20$ GeV, $|\eta| < 2.3$, and matching to a reconstructed $\tau_h$ candidate with $p_T > 18$ GeV. The different sources of improvement in performance of the algorithm with fixed strip size are shown separately for $\Delta \beta = 0.46$, $\Delta \beta = 0.46$ with $p_T^{\text{strip, outer}} < 0.1 p_T^{\tau_h}$ and for $\Delta \beta = 0.2$ with $p_T^{\text{strip, outer}} < 0.1 p_T^{\tau_h}$. The signal process is modelled using MC events for $H \rightarrow \tau\tau$ (for low-$p_T$ $\tau_h$) and $Z \rightarrow \tau\tau$ decays, with $m_Z = 2$ TeV (for high-$p_T$ $\tau_h$). The QCD multijet MC events are used as background, with jet $p_T$ values up to 100 and 1000 GeV, respectively, such that the $p_T$ coverage is similar to that in signal events. The improvement brought about by the dynamic strip reconstruction for high-$p_T$ $\tau$ leptons can be observed by comparing the two plots in figure 2.

At low-$p_T$ (figure 2, left), the performance for $\tau_h$ candidates for medium and tight WPs improves slightly. However, in the high-efficiency region, the misidentification probability starts to increase faster than the efficiency in the current algorithm. This is caused by choosing the working points of the algorithm through changes in the requirements on $I_{\eta\tau}$. To reach a higher efficiency, the requirement on $I_{\eta\tau}$ is relaxed, which in turn leads to an increase in the misidentification probability. However, the $p_T^{\text{strip, outer}}$ requirement prevents the efficiency from rising at a similar rate, leading thereby to the observed behaviour of the response in the high-efficiency region.

5.2.2 MVA-based discriminants

The MVA-based $\tau_h$ identification discriminants combine the isolation and other differential variables sensitive to the $\tau$ lifetime, to provide the best possible discrimination between $\tau_h$ decays and quark
or gluon jets. A classifier based on boosted decision trees (BDT) is used to achieve a reduction in the jet $\rightarrow \tau_h$ misidentification probability. The MVA identification method and the variables used as input to the BDT are discussed in ref. [24].

In addition to those discussed in ref. [24], the following variables are included in the classifier to improve its performance:

1. Differential variables such as $p_T^{\text{strip, outer}}$ in eq. (5.6), and $p_T$-weighted $\Delta R$, $\Delta\eta$, and $\Delta\phi$ (relative to the $\tau_h$ axis) of photons and electrons in strips within or outside of the signal cone;

2. $\tau$ lifetime information, based on the signed three-dimensional impact parameter of the leading track of the $\tau_h$ candidate and its significance (the impact parameter length divided by its uncertainty); and

3. multiplicity of photon and electron candidates with $p_T > 0.5$ GeV in the signal and isolation cones.

The charged and neutral-particle isolation sums and the $\Delta\beta$ correction, as defined in eq. (5.5), are used as separate variables in the BDT classifier, and correspond to the most powerful discriminating variables. Other significant variables are the two- and three-dimensional impact parameters of the leading track and their significances, as well as the flight length and its significance for the $\tau_h$ candidates decaying into three charged hadrons and a neutrino. The multiplicity of photon and electron candidates in the jet seeding the $\tau_h$ candidate is found to contribute to the decision of the BDT classifier at levels similar to those of the lifetime variables.

The BDT is trained using simulated $\tau_h$ candidates selected with $p_T > 20$ GeV and $|\eta| < 2.3$ in $Z/\gamma^* \rightarrow \tau\tau$, $H \rightarrow \tau\tau$, $Z' \rightarrow \tau\tau$, and $W' \rightarrow \tau\nu$ events (with the mass ranges of $H$, $Z'$, and $W'$...
detailed in section 3). The QCD multijet, $W +$jets, and $t\bar{t}$ events are used to model quark and gluon jets. These events are reweighted to provide identical two-dimensional distributions in $p_T$ and $\eta$ for $\tau_h$ candidates in signal and in background sources, which makes the MVA training insensitive to differences in $p_T$ and $\eta$ distributions of $\tau$ leptons and jets in the training samples.

The working points of the MVA-isolation discriminant, corresponding to different $\tau_h$ identification efficiencies, are defined through requirements on the BDT discriminant. For a given working point, the threshold on the BDT discriminant is adjusted as a function of $p_T$ of the $\tau_h$ candidate to ensure uniform efficiency over $p_T^{\tau_h}$. The working points for the reconstructed $\tau_h$ candidates are chosen to have isolation efficiencies between 40 and 90%, in steps of 10%, for the reconstructed $\tau_h$ candidates.

The expected jet $\rightarrow \tau_h$ misidentification probability is shown in figure 3, as a function of expected $\tau_h$ identification efficiency. It demonstrates a reduction in the misidentification probability by a factor of 2 for MVA-based discriminants, at efficiencies similar to those obtained using isolation-sum discriminants. We compare two sets of MVA-based discriminants that were trained using MC samples that correspond to different conditions during data taking. The working points of the MVA-based discriminants are shifted relative to each other, but follow the same performance curve. This confirms the stability of the MVA-based discriminants. The expected $\tau_h$ selection efficiencies and jet $\rightarrow \tau_h$ misidentification probabilities for low to medium $p_T$, for the most commonly used working point (tight) of the training in 2016 are 49% and 0.21%, respectively. For high $p_T$, the expected misidentification probability drops to 0.07%, while the $\tau_h$ selection efficiency remains constant, as desired.

![Figure 3. Misidentification probabilities for $\tau_h$ as a function of their identification efficiency, evaluated using $H \rightarrow \tau\tau$ (left), $Z'(2 \text{ TeV}) \rightarrow \tau\tau$ (right), and QCD multijet MC events. The MVA-based discriminants trained on their corresponding MC events are compared to each other, as well as to the isolation-sum discriminants. The points correspond to different working points of the discriminants. The three points for the isolation-sum discriminants from left to right correspond to the tight, medium, and loose WPs. Similarly, the six points of the MVA-based discriminants define the WP as very-very tight, very tight, tight, medium, loose, and very loose, respectively.](image-url)
Figure 4 shows the respective expected $\tau_h$ identification efficiency (left) and the misidentification probability (right), as a function of $p_T$ of the generated $\tau_h$ and of the reconstructed jet. The efficiency is computed from $Z \rightarrow \tau\tau$ events, while the expected jet $\mapsto \tau_h$ misidentification probability is computed for QCD multijet events with jet $p_T < 300\text{GeV}$.

![Efficiency and misidentification probability](image)

**Figure 4.** Efficiency of $\tau_h$ identification, estimated using simulated $Z/\gamma^* \rightarrow \tau\tau$ events (left), and the misidentification probability estimated using simulated QCD multijet events (right) are given, for the very loose, loose, medium, tight, very tight, and very-very tight WPs of the MVA-based $\tau_h$ isolation algorithm. The efficiency and misidentification probabilities are shown as a function of $p_T$ of the generated $\tau_h$ and of the reconstructed jet, respectively. Vertical bars (often smaller than the symbol size) correspond to the statistical uncertainties (the 68% Clopper-Pearson intervals [61]), while horizontal bars indicate the bin widths.

5.3 Discrimination of $\tau$ leptons against electrons

Isolated electrons have a high probability to be misidentified as $\tau_h$ candidate that decay to either $h^\pm$ or $h^\pm\pi^0$. In particular, electrons crossing the tracker material often emit bremsstrahlung photons mimicking neutral pions in their reconstruction. An improved version of the MVA electron discriminant used previously [24] is developed further to reduce the $e \mapsto \tau_h$ misidentification probability, while maintaining a high selection efficiency for genuine $\tau_h$ decays over a wide $p_T$ range. The variables used as input for the BDT are identical to the ones described in ref. [24], with the addition of the following photon-related variables:

1. the number of photons in any of the strips associated with the $\tau_h$ candidate;
2. the $p_T$-weighted root-mean-square of the distances in $\eta$ and $\phi$ between all photons included in any strip and the leading track of the $\tau_h$ candidate; and
3. the fraction of $\tau_h$ energy carried away by photons.

These variables are computed separately for the photons inside and outside of the $\tau_h$ signal cone to improve separation. The most sensitive variables are the fraction of energy carried by the photon candidates, the ratio of the energy deposited in the ECAL to the sum of energies deposited in the
ECAL and HCAL, the ratio of the deposited energy in the ECAL relative to the momentum of the leading charged hadron, the $m_{\tau}$, and the $p_T$ of the leading charged hadron.

The BDT is trained using the simulated events listed in section 3, which contain genuine $\tau$ leptons and electrons. Reconstructed $\tau_h$ candidates can be considered as signal or background, depending on whether they are matched to a $\tau$ decay or to an electron at the generator level. Different working points are defined according to the requirements on their BDT output and the efficiency for a genuine $\tau_h$ candidate to pass the working points of the discriminants. The expected efficiency of $\tau_h$ reconstruction and the $e \rightarrow \tau_h$ misidentification probability are presented in figure 5. Both are found to be approximately uniform over $p_T$, except for a dip at $\approx 45$ GeV, whose depth increases with the tightening of the selection criteria. This is because the MC events used to model the $e \rightarrow \tau_h$ misidentification in the training of the MVA-based electron discrimination algorithm have electron $p_T$ distributions that peak at $\approx 45$ GeV, since the sample is dominated by $Z/\gamma^* \rightarrow ee$ and $W \rightarrow e\nu$ events.

![Efficiencies and Misidentification Probabilities](image)

**Figure 5.** Efficiencies of $\tau_h$ identification estimated via simulated $Z/\gamma^* \rightarrow \tau\tau$ events (left), and the $e \rightarrow \tau_h$ misidentification probability estimated using simulated $Z/\gamma^* \rightarrow ee$ events (right) for the very loose, loose, medium, tight, and very tight WPs of the MVA-based electron discrimination algorithm. The efficiency is shown as a function of $p_T$ of the reconstructed $\tau_h$ candidate, while the misidentification probability is shown as a function of the generated electron $p_T$. The efficiency is calculated for $\tau_h$ candidates with a reconstructed decay mode that pass the loose WP of the isolation-sum discriminant, while the misidentification probability is calculated for generated electrons of $p_T > 20$ GeV and $|\eta| < 2.3$, excluding the less sensitive detector region of $1.46 < |\eta| < 1.56$ between the barrel and endcap ECAL regions. Vertical bars (often smaller than the symbol size) indicate the statistical uncertainties (the 68% Clopper-Pearson intervals), while horizontal bars indicate the bin widths.

### 5.4 Discrimination of $\tau$ leptons against muons

Muons have a high probability to be misreconstructed as $\tau_h$ objects in the $h^*$ decay mode. The discriminant against muons, developed previously [24], is based on vetoing $\tau_h$ candidates when signals in the muon detector are found near the $\tau_h$ direction. The two working points corresponding to different $\tau_h$ identification efficiencies and $\mu \rightarrow \tau_h$ misidentification rates are:
1. “against-\(\mu\) loose”: \(\tau_h\) candidates fail this working point when track segments in at least two muon detector planes are found to lie within a cone of size \(\Delta R = 0.3\) centered on the \(\tau_h\) direction, or when the energy deposited in the calorimeters, associated through the PF algorithm to the “leading” charged hadron of the \(\tau_h\) candidate, is <20% of its track momentum.

2. “against-\(\mu\) tight”: \(\tau_h\) candidates fail this working point when they fail condition (i), or when a hit is present in the CSC, DT, or RPC detectors located in the two outermost muon stations within a cone of size \(\Delta R = 0.3\) around the \(\tau_h\) direction.

The efficiency for \(\tau_h\) candidates from \(Z/\gamma^* \rightarrow \tau\tau\) events to pass the against-\(\mu\) discriminant selection requirements exceeds 99%. The \(\mu \rightarrow \tau_h\) misidentification probability, for muons in \(Z/\gamma^* \rightarrow \mu\mu\) events, is \(\approx 3.5 \times 10^{-3}\) and \(\approx 1.4 \times 10^{-3}\) for loose and tight WPs, respectively.

6 Reconstruction of highly boosted \(\tau\) lepton pairs

In events containing a (hypothetical) massive boson with large \(p_T\), e.g., a radion (R) decaying to a pair of Higgs bosons [62, 63], with at least one of these decaying to a pair of \(\tau\) leptons, the jets from the two \(\tau\) leptons would be emitted very close to each other, thereby forming a single jet. The performance of the HPS algorithm in such topologies is poor, as it was designed to reconstruct only one \(\tau_h\) per jet. A dedicated version of the HPS algorithm was therefore recently developed to reconstruct two \(\tau\) leptons with large momenta that typically originate from decays of large-momentum \(Z\) or Higgs bosons. This algorithm takes advantage of jet substructure techniques, as follows. A collection of “large-radius jets” is assembled from the PF candidates using the Cambridge-Aachen algorithm [64] with a distance parameter of 0.8 (CA8). Due to the large boosts, the emitted \(\tau\) lepton decay products are expected to be contained within the same CA8 jet, when its \(p_T\) exceeds 100 GeV. The algorithm proceeds by reversing the final step of the clustering algorithm for each given CA8 jet, to find two subjets \(s_{j1}\) and \(s_{j2}\) that can be expected to coincide with the two \(\tau\) leptons from the decay of the boosted massive boson. To reduce the misidentification of jets arising from QCD multijet events, \(s_{j1}\) and \(s_{j2}\) must satisfy the following additional restrictions:

1. the \(p_T\) of each subjet must be greater than 10 GeV, and

2. the mass of the heavier subjet must be less than 2/3 of the large-radius jet mass, where mass refers to the invariant mass of all jet constituents.

These requirements are obtained from an optimization of the reconstruction efficiency, while maintaining a reasonable misidentification probability. When these requirements cannot be met, the pair of subjets is discarded, and the procedure is repeated, treating the subjet with largest mass as the initial jet that is then split into two new subjets. If the algorithm is unable to find two subjets satisfying the above criteria within a given CA8 jet, no \(\tau_h\) reconstruction is performed from this CA8 jet, and the algorithm moves on to the next such jet. When two subjets satisfying the requirements are found, they are passed to the HPS algorithm as seeds. At this stage, the algorithm does not differentiate between subjets arising from hadronic or leptonic \(\tau\) decays. After reconstruction, the decay-mode criteria (section 5.1.2) and the MVA-based isolation discriminants (section 5.2.2) are applied to the
reconstructed $\tau_h$ candidate, taking into account just the PF candidates belonging to the subjet that seeds the $\tau_h$ in the reconstruction and the isolation calculations. The decay-mode criteria are relaxed relative to those used in the standard HPS algorithm by accepting $\tau_h$ candidates with two charged hadrons, and therefore an absolute charge different from unity. This relaxation recovers $\tau$ leptons decaying into three charged hadrons when one of the tracks is not reconstructed in the dense environment of a high-$p_T$ jet. If an electron or muon, reconstructed and identified through the usual techniques available for these leptons [21, 22], is found to be near ($\Delta R < 0.1$) to a $\tau_h$ candidate reconstructed from a subjet, the corresponding CA8 jet is considered to originate from a semileptonic $\tau$ lepton pair decay. Cases in which both $\tau$ leptons decay leptonically are not considered.

Figure 6 compares the efficiencies in standard reconstruction with that for highly boosted $\tau$ lepton pairs in simulated events of $R \rightarrow HH \rightarrow b\bar{b}\tau\tau$ decays in the $\tau_h\tau_h$ and $\tau_h\tau_\mu$ final states. In addition, the expected probability for large-radius jets to be misidentified as $\tau_h$ pairs is shown for simulated QCD multijet events. While the efficiency in $\tau_{\mu}\tau_h$ events is computed just for the $\tau_h$ candidate, it is computed once relative to one $\tau_h$ candidate and once relative to both $\tau_h$ candidates in $\tau_h\tau_h$ events. The misidentification probability is calculated in $\tau_h\tau_h$ final states for both $\tau_h$ candidates. The $\tau_h$ candidates are selected requiring $p_T > 20$ GeV and $|\eta| < 2.3$, using the very loose WP of the MVA-based isolation.

The algorithm used for highly boosted events provides a considerably higher efficiency than the standard HPS algorithm for $\tau$ lepton pairs with $p_T$ greater than $\approx 0.5$ TeV, with an expected increase in misidentification probability. Since at such high $p_T$, the contributions from background are highly suppressed, and the misidentification rate remains of the order of $10^{-4}$, this algorithm can be used for searches in this kinematic regime.

7 Identification of $\tau_h$ candidates in the high-level trigger

Several analyses are based on experimental signatures that include $\tau_h$ signals, and therefore, along with the offline reconstruction discussed in sections 5 and 6, we also employ dedicated $\tau_h$ identification algorithms in the trigger system, at both L1 and HLT.

The L1 system went through a series of upgrades [65] in 2015 and 2016, and it is now based on more powerful, fully-programmable FPGA processors and $\mu$TCA logic boards. This allows more sophisticated $\tau_h$ reconstruction and isolation algorithms at L1, the performance of which can be found in ref. [66].

The HLT system uses the full-granularity information of all CMS subdetectors, and runs a version of the CMS reconstruction that is slightly different than that used offline, as the HLT decision is made within 150 ms, on average, a factor of 100 faster than offline reconstruction. This is achieved using specialized, fast, or regional versions of reconstruction algorithms, and through implementation of multistep selection logic, designed to reduce the number of events processed by more complex, and therefore more time consuming subsequent steps. Both methods are exploited in the $\tau_h$ reconstruction at the HLT.

The $\tau_h$ HLT algorithm has three steps. The first step, referred to as Level 2 (L2), uses only the energy depositions in the calorimeter towers in regions around the L1 $\tau_h$ objects with $\Delta R < 0.8$. The depositions are clustered into narrow L2 $\tau_h$ jets using the anti-$k_T$ algorithm with a distance parameter of 0.2. The only selection criterion required at L2 is a $p_T$ threshold.
Figure 6. Reconstruction and identification efficiencies for the $\tau_h$ in the $\mu\tau_h$ (upper left) and $\tau_h\tau_h$ (upper right) final states, and for the $\tau_h\tau_h$ pair (lower left), as a function of the generated $p_T$ of the Higgs boson, and the probability for large-radius jets in QCD multijet events to be misidentified as $\tau_h\tau_h$ final states (lower right), as a function of the large-radius jet $p_T$. Vertical bars correspond to the statistical uncertainties (the 68% Clopper-Pearson intervals), while horizontal bars indicate the bin widths.

In the second step, known as Level 2.5 (L2.5), a simple form of charged-particle isolation is implemented, using just the information from the pixel detector. Tracks are reconstructed from hits in the pixel detector around the L2 $\tau_h$ jets (rectangular regions of $\Delta\eta\times\Delta\phi = 0.5\times0.5$), and used to form vertices. If no vertex is found, the $\tau_h$ jet is passed to the following step for more detailed scrutiny. If, on the other hand, at least one vertex is found, the one with highest $\sum p_T^2$ of its tracks is assumed to be the primary hard-scattering vertex in the event. Tracks originating from within $d_z < 0.1 \text{ cm}$ of the hard-scattering vertex, in an annulus of $0.15 < \Delta R < 0.4$ centered on the $\tau_h$ jet direction, and with at least three hits in the pixel detector, are used in the computation of the $\tau_h$ isolation.
jet isolation. An L2 $\tau_h$ jet is considered isolated if the scalar sum of the $p_T$ of the associated pixel tracks $\sum p_T^{\text{track}}$ is less than 1.85 GeV.

Finally, at Level 3 (L3), full track reconstruction, using both pixel and strip detectors, is executed using rectangular regions of size $\Delta \eta \times \Delta \phi = 0.5 \times 0.5$ around the L2 $\tau_h$ jets, followed by the PF reconstruction. Both components are tuned specifically for the fast processing at HLT, as discussed in ref. [55].

The L3 $\tau_h$ algorithm starts with jets clustered from PF particles by the anti-$k_T$ algorithm using a distance parameter of 0.4. First, photons, contained in a jet, within a fixed $\Delta \eta \times \Delta \phi$ area of 0.05×0.02 are clustered into the strips, and assigned the $\pi^0$ mass. A variable signal-cone size of $\Delta R_{\text{sig}}^{L3} = (3.6 \text{ GeV})/p_T^{\text{jet}}$, with $\Delta R_{\text{sig}}^{L3}$ limited to the range of 0.08–0.12, and an isolation cone of $\Delta R = 0.4$, are defined around the direction of the charged hadron in the jet with highest $p_T$. The L3 $\tau_h$ candidate is then constructed from the following constituents found within the signal cone: up to three charged hadrons that are ordered in decreasing $p_T$, and assumed to be charged pions, and all the available $\pi^0$ candidates. To recover possible tracking inefficiencies, neutral hadrons within a distance of $\Delta R = 0.1$ from the leading charged hadron are also considered as being part of the $\tau_h$ candidate. The vertex with smallest $d_z$ relative to the track of the leading charged hadron is considered as the vertex of the $\tau_h$ production. To maximize the HLT reconstruction efficiency, these identification criteria are chosen to be fairly inclusive, not requiring strict consistency with the $\tau_h$ decay modes, with the respective sizes of the signal and isolation cones chosen to be larger and smaller than the sizes of the corresponding cones in the offline algorithm.

Two types of isolations were defined for L3 $\tau_h$ candidates in 2016. First is the charged isolation ($\sum p_T^{\text{charged}}$), computed by summing the scalar $p_T$ of charged hadrons (other than those constituting the L3 $\tau_h$ candidate) with $d_z < 0.2$ cm relative to the $\tau_h$ vertex, located within the isolation cone; defining the loose, medium, and tight WPs through $\sum p_T^{\text{charged}}$ being smaller than 3.0, 2.0, and 1.5 GeV, respectively.

The second type is the combined isolation, $I_T^{L3}$, defined as

$$I_T^{L3} = \sum p_T^{\text{charged}} + 0.3 \max \left(0, \sum p_T^\gamma - p_T^{PU}\right),$$  

(7.1)

where $\sum p_T^\gamma$ is the sum of the scalar $p_T$ of photons within an annulus between the signal and isolation cones that do not belong to the signal strips, and $p_T^{PU}$ is the neutral contribution to the isolation from PU, estimated using the jet area method [56]. The respective loose, medium, and tight WPs of the combined isolation require $I_T^{L3}$ to be smaller than 3.0, 2.3, and 2.0 GeV.

The absolute isolation cutoff values (for both isolation types) are often relaxed by a few percent, depending on the trigger, as a function of $p_T^{\tau_h}$, starting at values of about twice the trigger threshold. This relaxation increases the reconstruction efficiency for genuine $\tau_h$ candidates, and is possible because of the number of misidentified $\tau_h$ candidates decreases with $p_T$, providing thereby a control of the trigger rates.

Finally, the scalar $p_T$ sum of photons that are included in the strips of the L3 $\tau_h$ candidate, but are located outside of its signal cone ($k_{\text{sig}}^{L3}$), is defined as for offline $\tau_h$ candidates in eq. (5.6). This variable was not used for $\tau_h$ triggers in 2016, but is included in triggers during data taking in 2017.

The $\tau_h$ reconstruction and identification algorithms described in this section are employed to define a set of triggers for data taking during 2016. The triggers and their performance are discussed in section 13.
8 Event selection and systematic uncertainties

This section describes the selection requirements employed to define event samples used in the following measurements of the performance of $\tau_h$ reconstruction and identification in data and simulation, as well as their related systematic uncertainties. Differences between data and simulated events in trigger, identification, and isolation efficiencies are taken into account through the reweighting of simulated events. In addition, the number of PU interactions in simulation is reweighted to match that measured in data.

8.1 The $Z/\gamma^* \rightarrow \tau\tau$ events

A sample of $Z/\gamma^*$ events decaying into $e\tau_h$ or $\mu\tau_h$ final states is selected by requiring at least one well-identified and isolated electron or muon, referred to as the “tag”, and one $\tau_h$ candidate that passes loose preselection criteria, which corresponds to the “probe”.

The events in the $e\tau_h$ final state are required to pass an isolated single-electron trigger with $p_T > 25$ GeV. Offline, the electron candidate is required to have $p_T > 26$ GeV and $|\eta| < 2.1$, pass the tight WP of the MVA-based electron identification (with an average efficiency of 80%) [21, 67], and have $I_{\text{rel}}^e < 0.1$, as defined in eq. (4.1). In the $\mu\tau_h$ final state, events are required to pass an isolated single-muon trigger with $p_T > 22$ GeV. Offline, the muon candidate is required to have $p_T > 23$ GeV and $|\eta| < 2.1$, pass the medium identification WP [22], and have $I_{\text{rel}}^\mu < 0.15$. The $\tau_h$ candidate is preselected to have $p_T > 20$ GeV, $|\eta| < 2.3$, no overlap with any global muon [22] with $p_T > 5$ GeV, to pass the against-lepton discriminant selection requirements defined in sections 5.3 and 5.4, and to have at least one charged hadron with $p_T > 5$ GeV. The $\tau_h$ and electron or muon are required to be separated by at least $\Delta R = 0.5$, and to carry opposite electric charges. If several $e\tau_h$ or $\mu\tau_h$ pairs in one event pass this set of selection criteria, the pair formed from the most isolated $\tau_h$ and the most isolated electron or muon is selected. The events are rejected if they contain an additional electron or muon passing relaxed selection criteria. The relaxed selection requires that an electron satisfies the very loose WP of the MVA-based identification (with an average efficiency of 95%), a muon has to be reconstructed as a global muon, and both the electron or muon must have $p_T > 10$ GeV and $I_{\text{rel}}^{e/\mu} < 0.3$. To reduce the W+jets background contribution, the transverse mass of the electron or muon and $p_T^{\text{miss}}$, $m_T = \sqrt{2p_T^{e/\mu}p_T^{\text{miss}}(1 - \cos \Delta \phi)}$, is required to be less than 40 GeV, where $\Delta \phi$ is the difference in azimuthal angle between the electron or muon $\vec{p}_T$ and $\vec{p}_T^{\text{miss}}$. In addition, a linear combination of the variables $P_{\xi}^{\text{miss}}$ and $P_{\xi}^{\text{vis}}$, originally developed by the CDF experiment [68], namely $D_\xi = P_{\xi}^{\text{miss}} / P_{\xi}^{\text{vis}} - 0.85 P_{\xi}^{\text{vis}}$, is used to benefit from the fact that in $Z/\gamma^* \rightarrow \tau\tau$ events the $p_T^{\text{miss}}$ from the neutrinos produced in $\tau$ decays typically forms a small angle with the visible $\tau_h$ decay products. The $D_\xi$ is required to be greater than $-25$ GeV.

8.2 The $\mu\tau_h$ final states in $t\bar{t}$ events

The $t\bar{t} \rightarrow \mu\tau_h$+jets events are selected in the same way as the $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\tau_h$ events, except for the requirements on $m_T$ and $D_\xi$, which are not applied. The events are also required to have at least one b-tagged jet to enrich the content in $t\bar{t}$ events.
8.3 The $Z/\gamma^* \rightarrow \mu\mu$ events to constrain the $Z/\gamma^* \rightarrow \ell\ell$ normalization

A high purity sample of $Z/\gamma^* \rightarrow \mu\mu$ events is selected to constrain the normalization of the Drell-Yan (DY, $q\bar{q} \rightarrow Z/\gamma^* \rightarrow \ell^+\ell^-$) events in the measurement of $\tau_\ell$ efficiency through the tag-and-probe method [69], described in detail in section 9.1. The events are required to have a pair of well-separated ($\Delta R > 0.5$), oppositely-charged muons. The leading (in $p_T$) muon is required to pass the same selection as used in the $\mu\tau_\ell$ final states of $Z/\gamma^*$ events. The subleading muon is required to pass the same selection as the leading muon, except for the $\eta$ requirement, which is relaxed to $|\eta| < 2.4$. The invariant mass of the dimuon pair is required to be within 60–120 GeV. Events are rejected if they contain an additional electron or muon passing the relaxed selection criteria.

8.4 Off-shell $W \rightarrow \tau\nu$ events

Here, we use events in which a virtual $W$ boson that decays into a $\tau_\ell$ and a $\nu$ is produced with small $p_T$ (and no accompanying hard jet). The $p_T$ of the $\tau_\ell$ and the $p_T^{\text{miss}}$ are expected to be well balanced in such events.

Events are required to pass a trigger where $p_T^{\text{miss}}_{\text{T,nos},0}$ and $H_T^{\text{miss}}_{\text{T,nos},0}$ are both greater than 110 GeV, with $p_T^{\text{miss}}_{\text{T,nos}}$ being the magnitude of $p_T^{\text{miss}}$ computed using all particles in an event except muons, and $H_T^{\text{miss}}_{\text{T,nos}}$ being the magnitude of $p_T^{\text{miss}}$ computed using jets with $p_T > 20$ GeV, reconstructed from all particles except muons. Offline, events are required to have one $\tau_\ell$ candidate with $p_T > 100$ GeV, and $p_T^{\text{miss}} > 120$ GeV. To ensure back-to-back topologies between the $\tau_\ell$ candidate and $p_T^{\text{miss}}$, we require $\Delta\phi(\tau_\ell, p_T^{\text{miss}}) > 2.8$ rad. The event is discarded if it has at least one jet with $p_T > 30$ GeV and $|\eta| < 4.7$, except the one corresponding to the $\tau_\ell$, or an additional electron or muon passing the relaxed selection criteria.

8.5 Off-shell $W \rightarrow \mu\nu$ events to constrain the $W \rightarrow \tau\nu$ normalization

This event sample is used to constrain the normalization of off-shell $W$ boson production for $m_W > 200$ GeV, used in the $\tau_\ell$ efficiency measurement, as described in section 9.3. Events are selected with an isolated single-muon trigger with $p_T > 22$ GeV and $|\eta| < 2.1$. Offline, the muon candidate must have $p_T > 120$ GeV and $|\eta| < 2.1$; it must also pass the medium identification WP, and have a relative isolation of less than 0.15. The event must also have $p_T^{\text{miss}} > 120$ GeV and $\Delta\phi(\mu, p_T^{\text{miss}}) > 2.8$ rad. The event is discarded if it has at least one jet with $p_T > 30$ GeV and $|\eta| < 4.7$, or an additional electron or muon passing the relaxed selection criteria.

8.6 Events from $W \rightarrow \mu\nu$+jet production

These events are triggered using a single isolated-muon trigger with $p_T > 24$ GeV and $|\eta| < 2.1$. Offline, we require one well-identified and isolated muon with $p_T > 25$ GeV. Events with additional electrons or muons passing the relaxed selection criteria are rejected. In addition, the transverse mass of the muon and $p_T^{\text{miss}}$ is required to be greater than 60 GeV, to suppress events with genuine $\tau_\ell$ candidates, in particular from $Z/\gamma^*$ bosons. Events should contain exactly one jet with $p_T > 20$ GeV and $|\eta| < 2.4$, and there should be no additional jets (in $|\eta| > 2.4$) with $p_T > 20$ GeV. To ensure that the $W$ boson is balanced in $p_T$ with the jet, the following selections are applied: $\Delta\phi(W, \text{jet}) > 2.4$ rad, and the ratio of jet $p_T$ and $W$ boson $p_T$ must be between 0.7 and 1.3, where the $p_T$ of the $W$ boson is reconstructed from the vector sum of muon $p_T$ and $p_T^{\text{miss}}$.
8.7 The $e\mu$ final states in $t\bar{t}$ events

These events are triggered using a single isolated-muon trigger with $p_T > 24$ GeV, and are required to have one well-identified and isolated electron and one well-identified and isolated muon both of $p_T > 26$ GeV and $|\eta| < 2.4$. Events with additional electrons or muons passing the relaxed selection criteria are rejected.

8.8 The $Z/\gamma^* \rightarrow ee$, $\mu\mu$ events for measuring the $e/\mu \rightarrow \tau_h$ misidentification probability

High-purity samples of $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ events are selected for measuring their respective $e \rightarrow \tau_h$ and $\mu \rightarrow \tau_h$ misidentification probabilities. Consequently, again, we require at least one well-identified, isolated electron or muon (tag) and one isolated $\tau_h$ candidate (probe).

The $Z/\gamma^* \rightarrow ee$ events are selected by requiring a single-electron trigger to have fired. Offline, the electron candidate must match the trigger object (within $|\Delta R| < 0.5$), have $p_T > 26$ GeV and $|\eta| < 2.1$, pass the most-restrictive electron-identification criteria, and have an $I_{rel}^e < 0.1$. The $Z/\gamma^* \rightarrow ee$ events are collected using a single isolated-muon trigger with $p_T > 24$ GeV. Offline, the muon candidate must match the trigger object (within $|\Delta R| < 0.5$), be selected with $p_T > 26$ GeV and $|\eta| < 2.1$, after passing medium muon-identification criteria, and $I_{rel}^\mu < 0.15$.

The $\tau_h$ candidate is required to satisfy $p_T > 20$ GeV and $|\eta| < 2.3$, be reconstructed in one of the decay modes $h^\pm$, $h^\pm\pi^0$, $h^\pm\pi^0\pi^0$, or $h^\pm h^\mp h^\pm$, and pass the tight WP of the MVA-based isolation discriminant described in section 5.2.2. It must also be separated from the electron or muon by $\Delta R > 0.5$, and have an electric charge opposite to that of the electron or muon. The $\tau_h$ candidate must pass the loose WP of the against-$\mu$ discriminant described in section 5.4 when selecting $Z/\gamma^* \rightarrow ee$ events. The purity of the sample is increased by requiring the invariant mass of the tag-and-probe pair to be between 60–120 or 70–120 GeV for $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ events, respectively.

The $W+\text{jets}$ and $t\bar{t}$ backgrounds are reduced by requiring the selected events to have $m_T$ (of the tag electron or muon and $\vec{p}_T^{\text{miss}}$) not exceeding 30 GeV.

8.9 Systematic uncertainties affecting all studied final states

The generic systematic uncertainties affecting most of the measurements presented in sections 9–12 are discussed in this section. Uncertainties concerning particular analyses are not covered here, but are discussed in their corresponding sections. The same is true for deviations in the values of the systematic uncertainties.

The uncertainty in the measured integrated luminosity is 2.5% [70], and affects the normalization of all processes modelled via MC simulation. The combination of trigger, identification, and isolation efficiencies for electrons and muons, measured using the tag-and-probe technique, result in normalization uncertainties of 2% that also affect the normalization of processes modelled in simulation. Uncertainties in the normalization of production cross sections [45–48, 50, 71, 72] or in the method used to extract the normalization of $t\bar{t}$ (3–10%), diboson (5–15%), and DY (2–4%) production, are also taken into account. Uncertainties in the $\tau_h$ energy scale, affecting the distributions in simulated events that depend on $E_{\tau_h}$, and range between 1.2% (as determined in section 12) and 3% for high-$p_T$ $\tau_h$ candidates. Furthermore, to account for statistical fluctuations caused by the limited number of simulated events, we use the “Barlow-Beeston light” approach [73, 74], which
assigns a single nuisance parameter per bin that rescales the total bin yield. Most of the analyses
discussed in the following sections correct the simulated $p_T$ distributions of the $Z/\gamma^*$ boson in DY
events and of the top quark in $t\bar{t}$ events to the spectra observed in data through measured weights.

This reweighting corrects only the differential distributions without changing their normalization.
Uncertainties in these weights are propagated through the analyses, where the downward changes by
one standard deviation are computed as a difference between the weighted distribution and the one
without weight, while the upward changes by one standard deviation are computed as a difference
between weighted distributions with nominal weight and with the square of that weight. Finally,
the uncertainty related to the PU distribution is estimated by changing the minimum-bias pp cross
section by ±5%.

A comprehensive overview of these uncertainties is given in table 2.

**Table 2.** Systematic uncertainties affecting the measurements described in sections 9–12. Given are the
source of the uncertainty and whether the distribution in question is affected.

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>Value</th>
<th>Affecting distribution?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated luminosity</td>
<td>2.5%</td>
<td>No</td>
</tr>
<tr>
<td>e trigger, identification, and isolation efficiency</td>
<td>2%</td>
<td>No</td>
</tr>
<tr>
<td>$\mu$ trigger, identification, and isolation efficiency</td>
<td>2%</td>
<td>No</td>
</tr>
<tr>
<td>DY normalization</td>
<td>2–4%</td>
<td>No</td>
</tr>
<tr>
<td>$t\bar{t}$ normalization</td>
<td>3–10%</td>
<td>No</td>
</tr>
<tr>
<td>Diboson normalization</td>
<td>5–15%</td>
<td>No</td>
</tr>
<tr>
<td>$\tau_h$ energy scale</td>
<td>1.2–3%</td>
<td>Yes</td>
</tr>
<tr>
<td>Limited number of events</td>
<td>Statistical uncertainty in individual bin</td>
<td>Yes</td>
</tr>
<tr>
<td>DY $p_T$</td>
<td>(Weight)$^2$ — no weight</td>
<td>Yes</td>
</tr>
<tr>
<td>$t\bar{t}$ $p_T$</td>
<td>(Weight)$^2$ — no weight</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of PU events</td>
<td>5%</td>
<td>Yes</td>
</tr>
</tbody>
</table>

## 9 Measurement of the $\tau_h$ identification efficiency

The measurements of $\tau_h$ reconstruction and identification efficiencies in data use approaches similar
to those of ref. [24], and provide data-to-simulation scale factors and their uncertainties that can
be used to correct the simulated predictions in analyses. The efficiency is measured in different
$p_T^{\tau_h}$ regions: small $p_T^{\tau_h}$ between 20 and $\approx$60 GeV, using the $\mu\tau_h$ final state of $Z/\gamma^* \rightarrow \tau\tau$ events,
as discussed in section 9.1; intermediate $p_T^{\tau_h}$ of up to $\approx$100 GeV, using the $\mu\tau_h$ final states in $t\bar{t}$
events, as discussed in section 9.2; and high $p_T^{\tau_h}$ of >100 GeV, using a selection of highly virtual W
bosons ($m_W > 200$ GeV) decaying into $\tau$ leptons, as presented in section 9.3. The data-to-simulation
scale factors obtained through these measurements are combined, as described in section 9.4, to
extrapolate to higher-$p_T^{\tau_h}$ regions not covered by these measurements. Finally, the identification
efficiency for $\tau_h$ candidates reconstructed using the algorithm dedicated to highly boosted $\tau$ lepton
pairs is measured using the tag-and-probe method, as described in section 9.5.
9.1 Using the tag-and-probe method in $Z/\gamma^*$ events

The $\tau_h$ identification efficiency for $p_T^{\tau_h}$ up to $\approx 60 \text{ GeV}$ is estimated in $\mu\tau_h$ final states of $Z/\gamma^*$ events, selected as described in section 8.1. The events are subdivided into passing (“pass” region) and failing (“fail” region) categories, depending on whether the $\tau_h$ candidate passes or fails the appropriate working point of the $\tau_h$ isolation discriminant. The data-to-simulation scale factor for the $\tau_h$ identification efficiency is extracted from a maximum likelihood fit of the invariant mass distribution of the reconstructed (visible) $\mu\tau_h$ system, referred to as $m_{\text{vis}}$. The expected SM contributions are fitted to the observed data simultaneously in both categories.

The predictions for SM processes contributing to the distribution in $m_{\text{vis}}$ consist of a signal sample of $Z/\gamma^* \to \tau\tau \to \mu\tau_h$ events, where the reconstructed $\tau_h$ candidate is required to be matched to the generated one, and a set of backgrounds. All background events, except for QCD multijet production, rely on simulated $m_{\text{vis}}$ distributions. Diboson, single top quark, and $t\bar{t}$ samples are normalized to their theoretical cross sections. A sample of dimuon events, as described in section 8.3, is used to constrain the normalization of the DY process, by using them simultaneously in the fit, along with the events in the passing and failing categories. The DY processes, other than the $Z/\gamma^* \to \tau\tau \to \mu\tau_h$ signal, where $\tau_h$ candidates from the misidentification of $e$, $\mu$, or jets, contribute to the background, and are denoted as “other DY”.

The normalization of the contribution from W+jets events is estimated using control samples in data. A data-to-simulation scale factor is estimated in a sample enriched in W+jets events, defined in a way similar to the signal sample, but without the $D_\xi$ requirement having been applied, and with $m_T > 80 \text{ GeV}$, where small contributions from other processes are subtracted from data, based on their estimated cross sections. The scale factor is then applied to the simulation of the W+jets events in the low-$m_T$ signal sample.

The distribution and normalization of the QCD multijet background is estimated from control samples in data. The distribution is extracted from a sample selected using the nominal selection criteria discussed previously, but requiring the $\mu$ and $\tau_h$ candidates to have the same-sign (SS) electric charge. All other processes contributing to this sample are estimated using the procedures detailed above, and are subtracted from the data. The normalization is controlled using the ratio of events found in two separate control samples requiring same- and opposite-sign (OS) charge for the $\mu$ and $\tau_h$ candidates, respectively. Otherwise, both samples are defined in ways similar to that of the signal sample, but with an inverted muon isolation criterion.

The following uncertainties are considered in addition to the ones outlined in section 8.9, that is, uncertainties in the W+jets background normalization that arise from a possible difference between the low- and high-$m_T$ regions and from the uncertainties in $p_T^{\text{miss}}$, which are used in computing $m_T$. The uncertainty in the yield of W+jets events is estimated to be about 10%. The uncertainty in the OS/SS scale factor, used in the estimation of the QCD multijet background is $\approx 5\%$, which is mostly due to the limited number of events in the OS and SS control regions. The normalization of the DY process is extracted from the dimuon control region. An extrapolation uncertainty of 2% is used for the $\mu\tau_h$ sample to account for the differences in lepton kinematics (mostly in $p_T$).

The results obtained for different working points of the MVA-based discriminant with $\Delta R = 0.5$ are shown in table 3. An uncertainty of 3.9% is added in quadrature to the one returned by the fit, to account for the uncertainty associated with the track reconstruction efficiency [26]. The scale
factors obtained for different working points of the isolation sum discriminants are found to be close to 90%, with uncertainties of 5%, and the scale factors obtained for the MVA-based discriminants trained using 2016 simulations as well as for $\Delta R = 0.3$, are found to be compatible with those presented in table 3. The measured scale factors vary from 0.92 to 0.99, depending on the working point, with uncertainties of about 5%. The fitted distributions that maximize the likelihood for the tight WP of the MVA-based isolation are shown in figure 7. The scale factors are also measured in different ranges of $p_T^{\tau_h}$ for the tight WP of the MVA-based isolation discriminant with $\Delta R = 0.5$, and enter the extrapolation to high $p_T^{\tau_h}$, as discussed in section 9.4.

Table 3. Data-to-simulation scale factors for different MVA-based isolation working points with $\Delta R = 0.5$, measured using $Z/\gamma^*$ events. An uncertainty of 3.9% is added in quadrature to the uncertainty returned by the fit to account for the uncertainty in track reconstruction efficiency.

<table>
<thead>
<tr>
<th>Working point</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very loose</td>
<td>0.99 ± 0.05</td>
</tr>
<tr>
<td>Loose</td>
<td>0.98 ± 0.05</td>
</tr>
<tr>
<td>Medium</td>
<td>0.97 ± 0.05</td>
</tr>
<tr>
<td>Tight</td>
<td>0.95 ± 0.05</td>
</tr>
<tr>
<td>Very tight</td>
<td>0.92 ± 0.05</td>
</tr>
<tr>
<td>Very-very tight</td>
<td>0.93 ± 0.05</td>
</tr>
</tbody>
</table>

Figure 7. The fitted distribution in $m_{\text{vis}}$ in the passing (left) and failing (right) categories for the tight WP of the MVA-based isolation. The electroweak background includes contributions from $W$+jets (dominating), diboson, and single top quark events. Vertical bars correspond to the statistical uncertainties in the data points (68% frequentist confidence intervals), while shaded bands to the quadratic sum of the fitted statistical and systematic uncertainties.

The efficiency for $\tau_h$ candidates to pass the working points of the discriminants used to reject electrons and muons, described in sections 5.3 and 5.4, respectively, are also measured in the $\mu\tau_h$ final states of $Z/\gamma^* \rightarrow \tau\tau$ events, which are selected as described above. The $\tau_h$ candidates are
required to have $p_T > 20$ GeV, $|\eta| < 2.3$, and to pass the tight WP of the MVA-based $\tau_h$ isolation discriminant. The events are again subdivided into passing and failing categories, depending on whether the $\tau_h$ candidate passes or fails the appropriate working points of the discriminants used against electrons or muons. The data-to-simulation scale factor is obtained from a maximum likelihood fit to the $m_{\text{vis}}$ distribution. The scale factors are compatible with unity to within the uncertainty in the measurements that range between 1 and 3%.

### 9.2 Using $t\bar{t}$ events

A sample of $t\bar{t}$ events with a muon and a $\tau_h$ in the final state is used to measure the $\tau_h$ identification efficiency for $p_T^{\tau_h}$ up to 100 GeV. The selection requirements are described in section 8.2. The selected $\tau_h$ candidate must be accepted using the appropriate working point of the $\tau_h$ isolation discriminant. The distribution in $m_T$ of the muon and $\vec{p}_T^{\text{miss}}$ is used to determine the data-to-simulation scale factors.

Contributions to $m_T$ distribution from $Z/\gamma^* \rightarrow \tau\tau$, single top quark, diboson, and $W+$jets events are modelled using simulations normalized to theoretical cross sections. Background from QCD multijet production is determined as described in section 9.1. The major background contribution is from $t\bar{t}$ events where a jet is misidentified as a $\tau_h$ candidate. The distribution is taken from simulation and a dedicated sample of events is selected to constrain the normalization of this background, as well as the probability of a jet to be misidentified as a $\tau_h$ candidate. Events have to pass the same criteria as discussed in section 8.2, but must also contain an additional isolated electron of electric charge opposite to that of the selected muon. This selects the $e\mu$ final state of $t\bar{t}$ events with an additional jet which can be misidentified as a $\tau_h$ candidate. These $e\mu$ events are then subdivided into passing and failing categories, based on whether the requirements imposed on the $\tau_h$ candidate are met in the $\tau_h$ isolation discriminant. A simultaneous likelihood fit is performed to the $m_T$ distribution in all three samples, constraining thereby the $t\bar{t}$ contribution and the probability for jets to be identified as $\tau_h$ candidates, as well as measuring the efficiency of the $\tau_h$ identification relative to that expected in simulation.

The systematic uncertainties are similar to those listed in section 8.9, except for additional uncertainties related to the $b$ tagging performance (3% effect on the normalization), and the cross section for $Z/\gamma^*+\text{jet}$ process (30%), given that the $Z/\gamma^*+b$ jet cross section is not well measured. A 3.9% uncertainty in the track reconstruction efficiency is added to the signal processes. The uncertainty in the jet $\rightarrow \tau_h$ misidentification probability is correlated between the signal and the control sample, where the $\tau_h$ candidate passes the identification requirement. The $e\mu$ failing category is used to further constrain both the normalization for $t\bar{t}$ production as well as the uncertainty in $b$ tagging performance. Figure 8 shows the fitted distributions in $m_T$ for the tight WP of the MVA-based isolation.

The measurement is repeated for different isolation working points of the MVA-based discriminant, as well as for the tight WP in different regions of $p_T^{\tau_h}$, and individually for each reconstructed decay mode. Although the mean value of the scale factor in the $h^+h^-\tau^\pm\tau^\mp$ decay mode is slightly below those of the other decay modes, no significant differences are observed between the three decay modes. The measured scale factors in different $p_T^{\tau_h}$ regions enter the extrapolation as outlined in section 9.4, and table 4 summarizes the results for the working points of the MVA-based isolation discriminants. The scale factors measured from the inclusive $t\bar{t}$ events are slightly lower than those from $Z/\gamma^* \rightarrow \tau\tau$. This is because the jet $\rightarrow \tau_h$ misidentification probability is slightly higher.
Figure 8. Fitted distributions for the signal (upper), $e\mu$ passing (lower left), and the $e\mu$ failing (lower right) categories, using the $m_T$ for the $p_T^D$ and $p_T^{\text{miss}}$ vectors as observables for the tight WP of the MVA-based isolation with $p_T^D$ between 30 and 40 GeV. The electroweak background includes contributions from W+jets (dominating), diboson, and single top quark events. Vertical bars correspond to the statistical uncertainties in the data points (68% frequentist confidence intervals), while the shaded bands reflect the quadratic sum of the statistical and systematic uncertainties after the fit.

Table 4. Data-to-simulation scale factors for different MVA-based isolation working points obtained from $\bar{t}t$ events.

<table>
<thead>
<tr>
<th>Working point</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very loose</td>
<td>0.99 ± 0.07</td>
</tr>
<tr>
<td>Loose</td>
<td>0.94 ± 0.07</td>
</tr>
<tr>
<td>Medium</td>
<td>0.91 ± 0.07</td>
</tr>
<tr>
<td>Tight</td>
<td>0.92 ± 0.06</td>
</tr>
<tr>
<td>Very tight</td>
<td>0.89 ± 0.06</td>
</tr>
<tr>
<td>Very-very tight</td>
<td>0.86 ± 0.06</td>
</tr>
</tbody>
</table>
in simulation than in data, causing the $\tau_h$ identification efficiency scale factor to be pulled down towards lower values, where the distributions of the $t\bar{t}$ events with genuine $\tau_h$ and the misidentified jet $\rightarrow \tau_h$ candidates become similar. However, this is mitigated for the measurement in bins of $p_T^{\tau_h}$, by constraining the normalization of the $t\bar{t}$ background with a jet misidentified as a $\tau_h$ candidate, using the $e\mu$ passing sample, as discussed above.

9.3 Using off-shell $W \rightarrow \tau\nu$ events

The identification efficiency for $\tau_h$ leptons with $p_T > 100$ GeV is measured using a sample of events in which a highly virtual $W$ boson ($m_W > 200$ GeV) is produced at small $p_T$ (and often without an accompanying hard jet), and decays into a $\tau$ lepton and $\nu_\tau$. The signature for such events consists of a single $\tau_h$ decay and $p_T^{miss}$ balanced by the $p_T^{\tau_h}$.

The selection requirements for the $W \rightarrow \tau\nu$ sample are described in section 8.4. A large fraction of events selected in this channel originate from processes where a jet is misidentified as a $\tau_h$ candidate. The main processes contributing to this kind of background are QCD multijet, $Z/\gamma^* \rightarrow \nu\bar{\nu} +$jets, and $W \rightarrow \ell\nu +$jets events.

The background from events where a jet is misidentified as a $\tau_h$ candidate is estimated using a control sample obtained by applying the same set of requirements as used in the selection of the $W \rightarrow \tau\nu$ events, except for the $\tau_h$ isolation criterion, which is inverted. Events in this control sample are then extrapolated to the signal region using the ratio of probabilities for a jet to pass to that to fail the $\tau_h$ isolation. The $W \rightarrow \mu\nu +1$ jet and QCD dijet events are utilized to estimate the extrapolation factor. The method is verified with simulated samples of $W \rightarrow \ell\nu +$jets and $Z/\gamma^* \rightarrow \nu\bar{\nu} +$jets events.

The study shows that the set of requirements outlined in section 8.4, selects $W \rightarrow \tau\nu$ events with an invariant mass of the $\tau\nu$ pair $m_{\tau\nu} \equiv m_W > 200$ GeV. A dedicated auxiliary sample of $W \rightarrow \mu\nu$ events is used to constrain the normalization of virtual $W$ boson production with $m_W > 200$ GeV. The $W \rightarrow \mu\nu$ events are selected as described in section 8.5, and verified using MC simulation that the phase space covered by the $W \rightarrow \mu\nu$ and $W \rightarrow \tau\nu$ samples tend to largely overlap.

The signal is extracted using a simultaneous maximum likelihood fit to the $m_T$ (of the $p_T^{\tau_h}/\mu$ and $p_T^{miss}$) distribution for both the $W \rightarrow \tau\nu$ signal and $W \rightarrow \mu\nu$ control samples. This procedure minimizes the uncertainties related to the normalization of $W$ boson events. The fit is performed using two freely floating parameters:

1. the scale factor in the $\tau_h$ identification efficiency, i.e., the ratio of the measured value of the $\tau_h$ identification efficiency to the value predicted by simulation, and

2. the normalization for $W$ production with $m_W > 200$ GeV, relative to the theoretical prediction ($r_W$).

In addition to the uncertainties listed in section 8.9, the following systematic uncertainties are also taken into account in the fit. An uncertainty of 1% in the momentum scale of the muon that also alters the differential distributions. The energy scale of the $p_T^{miss}$ is taken into account in propagating the uncertainty in the jet energy scale, as well as in the scale of the unclustered energy depositions. Uncertainties in the extrapolation factor used in the estimation of background from jets misidentified as $\tau_h$ is also taken into account. The backgrounds with genuine $\tau$ leptons in $W \rightarrow \tau\nu$ events are dominated by diboson events, which are estimated via MC simulation. The normalization of the diboson...
background is verified in dedicated control regions, indicating discrepancies of up to 30%. An uncertainty of 30% is therefore used in the normalization of backgrounds containing genuine $\tau$ leptons.

Figure 9 shows the fitted $m_T$ distributions for the $W \rightarrow \tau \nu$ signal and $W \rightarrow \mu \nu$ control samples. The scale factor in the $\tau_h$ identification efficiency, the parameter $r_W$, and the correlation coefficient between the two quantities obtained from the fits, are detailed in Table 5 for different working points of the MVA-based $\tau_h$ isolation discriminants. The data-to-simulation scale factors range between 0.89 for the very tight WP and 0.96 for the loose WP. The fitted value of the W boson production cross section for $m_W > 200$ GeV is consistent with theoretical predictions. The W boson sample normalization factor is anticorrelated with the scale factor for $\tau_h$ identification efficiency, as an increase in the W boson yield is compensated in the fit by a reduction in the scale factor. The correlation between the scale factor and $r_W$ increases with tighter $\tau_h$ isolation, as expected, due to an increase in the purity of the signal region.

Table 5. The scale factor in the $\tau_h$ identification efficiency, the normalization of W boson production with $m_W > 200$ GeV, $r_W$, and the correlation coefficient between the two quantities obtained from the fit, measured for MVA-based discriminants using $\Delta R = 0.5$ in $W \rightarrow \tau \nu$ events.

<table>
<thead>
<tr>
<th>Working point</th>
<th>Scale factor</th>
<th>$r_W$</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loose</td>
<td>0.96 ± 0.08</td>
<td>1.03 ± 0.06</td>
<td>−0.34</td>
</tr>
<tr>
<td>Medium</td>
<td>0.93 ± 0.07</td>
<td>1.02 ± 0.07</td>
<td>−0.44</td>
</tr>
<tr>
<td>Tight</td>
<td>0.91 ± 0.07</td>
<td>1.02 ± 0.07</td>
<td>−0.46</td>
</tr>
<tr>
<td>Very tight</td>
<td>0.89 ± 0.07</td>
<td>1.02 ± 0.06</td>
<td>−0.47</td>
</tr>
</tbody>
</table>

We also measure the $\tau_h$ identification efficiencies in bins of $p_T^{\tau_h}$, with the data-to-simulation scale factors extracted in a simultaneous fit to the $m_T$ distribution in four signal samples, corresponding
to four bins of \( p_{T}^{h} \), and of \( p_{T}^{l} \) in the \( W \rightarrow \mu \nu \) control sample. The results enter in the extrapolation of the scale factor to high \( p_{T}^{h} \), as discussed in section 9.4.

### 9.4 Extrapolation of the \( \tau_{h} \) identification efficiency to large \( p_{T}^{h} \)

To extrapolate the scale factors for the \( \tau_{h} \) identification efficiency to high \( p_{T}^{h} \), a fit is performed to the values obtained in sections 9.1, 9.2, and 9.3, as a function of \( p_{T}^{h} \). These measurements cover a \( p_{T}^{h} \) range between 20 and \( \approx 300 \) GeV, with the mean value in each \( p_{T}^{h} \) bin used as a representative number for that bin. Fits to a zero- (constant) and first-order polynomial are performed, without considering the uncertainty in track reconstruction efficiency, as it is correlated among the individual measurements. Nevertheless, it is found to contribute very little to the overall uncertainty, with the exception of measurements at low \( p_{T}^{h} \), where other uncertainties are small because of the large number of events and the high purity of the event samples. Despite having other possible correlations between \( p_{T}^{h} \) bins in a single measurement, or between different measurements, all measurements entering the fit are assumed to be uncorrelated.

The fit to a first-order polynomial provides a smaller goodness-of-fit per degree of freedom, \( \chi^{2}/\text{dof} \), than that to a constant, indicating that the scale factor for \( \tau_{h} \) identification efficiency may decrease with \( p_{T}^{h} \); but, given that the slope of the fitted first-order polynomial barely deviates from zero (by only about one standard deviation), the scale factor is compatible with being constant. As there are no indications that components of \( \tau_{h} \) reconstruction or identification behave abnormally at high \( p_{T}^{h} \), a constant scale factor with an asymmetric uncertainty that increases with \( p_{T}^{h} \) is defined by adding in quadrature the uncertainty in the fit to a constant, and the difference between the fit to a first-order polynomial and to a constant for the downward deviation. In addition, this also takes into account the uncertainty in the efficiency of track reconstruction, yielding the total (asymmetric) uncertainty of \(+5\% \times p_{T}^{B} \) (TeV) and \(-35\% \times p_{T}^{B} \) (TeV). The fit to a constant using the combined uncertainty is shown in figure 10.

### 9.5 Using the tag-and-probe method in \( Z/\gamma^{*} \) events for highly boosted \( \tau \) lepton pairs

The identification efficiency for highly boosted \( \tau \) lepton pairs in \( \tau_{h} \) final states is measured using the same tag-and-probe method as described in section 9.1. The selection is optimized to have a pure sample of \( \tau \) leptons from the decay of high-\( p_{T} \) Z bosons, where one \( \tau \) lepton decays leptonically and the other one into hadrons and a neutrino. As the trigger thresholds for nonisolated leptons are very high, too few events are available to reliably measure the identification efficiency for very high \( p_{T} \) \( \tau \) lepton pairs. Single-isolated-lepton triggers with lower thresholds are used therefore to select \( e\tau_{h} \) and \( \mu\tau_{h} \) events. However, events in which a \( \tau_{h} \) is within the isolation area around a triggering lepton (\( \Delta R < 0.4 \)) are not accessible in this measurement.

The selection requires one isolated electron or muon fulfilling tight identification criteria, and satisfying \( p_{T} > 40 \) or \( >26 \) GeV, respectively. Furthermore, as discussed in section 6, at least one \( \tau_{h} \) candidate must be reconstructed with \( p_{T} > 20 \) GeV and \( |\eta| < 2.3 \), in compliance with relaxed decay mode criteria. The \( \Delta R \) between the selected lepton and \( \tau_{h} \) candidate must be between 0.4 and 0.8, and the \( m_{T} \) of the \( \vec{p}_{T}^{e} \) and \( \vec{p}_{T}^{\text{miss}} \) system must be \( < 40 \) GeV. Moreover, \( p_{T}^{\text{miss}} \) must exceed 75 GeV, the scalar \( p_{T} \) sum of all measured particles has to be greater than 200 GeV, and there cannot be any identified b jets in the event. If more than one \( e\tau_{h} \) or \( \mu\tau_{h} \) pair is present, the one with the largest \( p_{T} \) is chosen for further analysis.
The contribution from DY events is modelled using MC simulation. It is split into the signal contribution by matching the reconstructed leptons to those generated and those contributing via misidentified Z boson decays. The distributions of the backgrounds from W+jets and t̄t̄ production are also modelled using simulation, but their normalizations are obtained from dedicated control data samples. The control sample for W+jets production is defined by inverting the requirement on \( m_T \). The control sample for t̄t̄ production is established by demanding at least one b-tagged jet.

The background from QCD multijet production is estimated from a sample selected in the same way as the signal, except for the requirement on \( p_T^{\text{miss}} \), which is inverted to \( p_T^{\text{miss}} < 75 \text{ GeV} \). Contributions from other processes are subtracted based on simulation. The extrapolation factor from the sample with an inverted \( p_T^{\text{miss}} \) requirement to the signal region is obtained from the ratio of events in two other control samples, where the \( \Delta R \) between the lepton and the \( \tau_h \) candidate is between 0.8 and 2.0, one which uses the nominal and the other an inverted \( p_T^{\text{miss}} \) requirement, respectively. Contributions from other processes are also subtracted from data using MC simulation in these two control regions.

The systematic uncertainties discussed in section 8.9 are taken into account in the procedure, as are the additional uncertainties in the estimation of the QCD multijet background, which are dominated by the limited number of events in the control samples. Finally, the uncertainties in the normalization of background from t̄t̄ and W+jets production are determined from their respective control samples, and amount to 3 and 13%, respectively.

The data-to-simulation scale factors are evaluated in the same way as outlined in section 9.1. The passing and failing events are defined by requiring the \( \tau_h \) to pass or fail a given working point of the MVA-based isolation discriminant. The scale factors for the six MVA-based working points are shown in table 6. The values are compatible with unity, as well as with the scale factors obtained...
Table 6. Data-to-simulation scale factors for different working points of the MVA-based isolation discriminant, using highly boosted $Z/\gamma^*$ events decaying to $\tau$ lepton pairs.

<table>
<thead>
<tr>
<th>Working point</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very loose</td>
<td>0.97±0.09</td>
</tr>
<tr>
<td>Loose</td>
<td>0.99±0.09</td>
</tr>
<tr>
<td>Medium</td>
<td>0.98±0.09</td>
</tr>
<tr>
<td>Tight</td>
<td>0.96±0.08</td>
</tr>
<tr>
<td>Very tight</td>
<td>0.95±0.09</td>
</tr>
<tr>
<td>Very-very tight</td>
<td>0.90±0.08</td>
</tr>
</tbody>
</table>

Figure 11. Fitted distributions to the passing (left) and failing (right) events for $\tau_h$ from highly boosted $\tau$ lepton pairs that pass the medium WP of the MVA-based isolation discriminant. The electroweak background includes contributions from $W$+jets (dominating), diboson, and single top quark events. Vertical bars correspond to the statistical uncertainties in the data points (68% frequentist confidence intervals), while the shaded bands provide the quadratic sum of the statistical and systematic uncertainties after the fit.

10 Measurement of the jet $\mapsto \tau_h$ misidentification probability

10.1 Using $W \to \mu\nu$+jet events

The probability to misidentify a quark or gluon jet as a $\tau_h$ candidate is measured as a function of jet $p_T$ and $\eta$ in a sample of $W \to \mu\nu$+jet events, selected as described in section 8.6. In addition to $p_T^{\text{jet}}$ and $\eta^{\text{jet}}$, the misidentification probability also depends on parton flavour, as well as whether the parton initiating the jet and the reconstructed $\tau_h$ have the same or opposite charge. These factors cause differences of up to a factor of four between misidentification probabilities for c quark and gluon jets, and up to a factor of two for whether the initiating parton has the same or opposite charge...
as the $\tau_h$ candidate. This means that the misidentification probabilities given in this section are indicative, in that they are mainly valid for $W \rightarrow \mu\nu+\text{jet}$ events, which contain a large fraction of light-quark jets, and therefore have a relatively high misidentification probability.

The misidentification probability is given by the ratio of the number of jets that are identified as $\tau_h$ candidates with $p_T > 20$ GeV, $|\eta| < 2.3$, and passing any one of the working points of the discriminants described in section 5.2, to the total number of jets with $p_T > 20$ GeV and $|\eta| < 2.3$. It should be recognized that $p_T^{\text{jet}}$ differs from $p_T^{\tau_h}$ because the four-momentum of the jet is computed by summing the momenta of all its constituents, while the $\tau_h$ four-momentum is computed only from the charged hadrons and photons used in the reconstruction of the specified decay mode of the $\tau_h$ candidate. For $p_T^{\text{jet}} < 300$ GeV, the $p_T^{\tau_h}$ constitutes on average only 40% of the jet $p_T$. Furthermore, $p_T^{\text{jet}}$ is subject to additional jet energy corrections, whereas $p_T^{\tau_h}$ is not.

In the measurement of the misidentification probability, backgrounds with genuine $\tau_h$ are subtracted, based on the expectations from simulated events. The fraction of events with genuine $\tau_h$ candidates in the sample passing the $\tau_h$ identification criteria is well below 10% for $\tau_h$ with $p_T < 100$ GeV, but reaches up to 50% for $p_T \approx 300$ GeV. Furthermore, backgrounds with prompt electrons and muons giving rise to $\tau_h$ candidates are also subtracted based on expectations from simulated events. To reject events from $Z/\gamma^* \rightarrow \mu\mu$ production, the loose WP of the against-$\mu$ discriminant described in section 5.4 is applied to the reconstructed $\tau_h$ candidates.

The subtraction of backgrounds containing genuine $\tau_h$ is subject to an uncertainty of 30%, leading to an uncertainty of up to 15% in the jet $\rightarrow \tau_h$ misidentification probability. Because of threshold effects, the jet energy scale also leads to a significant uncertainty, especially in the lowest bin of $p_T^{\text{jet}}$. Additional uncertainties are considered for probabilities with which electrons are reconstructed as $\tau_h$ candidates (with $\approx 100\%$ relative values), and with which muons are reconstructed as $\tau_h$ candidates that pass the loose WP of the against-$\mu$ discriminant (at 50%). These lead to uncertainties in the measured misidentification probabilities of at most a few percent.

The observed and simulated jet $\rightarrow \tau_h$ misidentification probabilities for the loose, medium, and tight WPs of the MVA-based isolation discriminant are shown in figure 12, as a function of $p_T^{\text{jet}}$ and $\eta^{\text{jet}}$. The probabilities are observed to be almost constant as a function of $\eta^{\text{jet}}$, while they decrease monotonically with increasing $p_T^{\text{jet}}$ from $\approx 40$ GeV, as the absolute isolation increases for quark- and gluon-initiated jets with increasing jet $p_T$. The values of the misidentification probability as a function of $p_T^{\text{jet}}$ range between 2.0 and 0.1% for the loose WP of the MVA-based isolation discriminant, and between 1.0 and less than 0.1% for the tight WP. The observed probabilities show a difference of 10–20% relative to expectations from MC simulation. This difference is well within the range of the misidentification probabilities obtained under variations of the parton shower models and underlying-event tunes, and reflects precision of modelling untypical, narrow and low multiplicity, quark and gluon jets being able to pass $\tau_h$ identification criteria.

10.2 Using $e\mu+$jets events

The probability to misidentify quark and gluon jets as $\tau_h$ candidates is also measured in the $e\mu$ final state of $t\bar{t}$ events using the same methodology and uncertainties outlined in section 10.1. The events are selected as described in section 8.7, with the largest contributions being from $t\bar{t}$ and single top quark events, where the misidentified $\tau_h$ candidates are dominated by $b$ quark jets. The contribution from other processes is $<10\%$. The observed and simulated jet $\rightarrow \tau_h$ misidentification
Figure 12. Probabilities for quark and gluon jets in \( W \)+jet events to pass the loose (uppermost), medium (middle), and tight (lowest) WPs of the MVA-based isolation discriminant as a function of \( p_T^{\text{jet}} \) (left) and \( \eta^{\text{jet}} \) (right). The misidentification probabilities in data are compared to expectations from simulation. The vertical bars in the simulated and observed misidentification probabilities include statistical uncertainties from the limited event count in both data and simulated samples, including the background subtraction. The shaded bands contain the systematic uncertainties related to background subtraction and the jet energy scale.
probabilities for the loose, medium, and tight WPs of the MVA-based isolation discriminant are shown in figure 13, as a function of $p_{\text{jet}}^T$ and $\eta_{\text{jet}}$. The observed probabilities show a 10–20% difference relative to expectations from simulation, except in a few $\eta_{\text{jet}}$ bins where the differences are as large as 50%. The jet $\tau_h$ misidentification probabilities in $e\mu$+jets events are found to be smaller than those for $W$+jet events because of the larger fraction of $b$ quark jets. The $b$ quark jets are typically less collimated than the light-quark jets, providing thereby smaller probabilities to pass the $\tau_h$ isolation discriminant selection requirements.

11 The $e/\mu \rightarrow \tau_h$ misidentification probability

11.1 Measurement of the $e \rightarrow \tau_h$ probability

The $e \rightarrow \tau_h$ misidentification probability is obtained from data using a tag-and-probe method in $Z/\gamma^* \rightarrow ee$ events selected as described in section 8.8. Depending on whether the probe passes or fails a given working point of the against-$e$ discriminant, the event enters the passing or failing category, respectively. The $e \rightarrow \tau_h$ misidentification rate is then measured in a simultaneous fit to the number of $Z/\gamma^* \rightarrow ee$ events in both categories. The $m_{\text{vis}}$ distribution in the range $60 < m_{\text{vis}} < 120$ GeV is used in the passing category, obtained from the templates for $Z/\gamma^* \rightarrow ee$ signal and for the $Z/\gamma^* \rightarrow \tau\tau$, $W$+jets, $t\bar{t}$, single top quark, diboson ($WW$, $WZ$, $ZZ$), and QCD multijet backgrounds. In the failing category, the total number of events in the same range of $m_{\text{vis}}$ is used to constrain the normalization of the $Z/\gamma^* \rightarrow ee$ process.

The differential templates for signal and all background distributions, except for QCD multijet, are taken from MC simulation. The normalization is performed according to the cross section for the specific sample of events, with the exception of the $W$+jets background, which is obtained from data, using an enriched sample of $W$+jets events with $m_T > 70$ GeV. The scale factor between the sideband and the signal region is extracted from simulation. The differential distribution and normalization of the QCD multijet background is obtained from data in a control sample where the tag and the probe are of SS. The contributions from all other backgrounds are estimated using simulation, and are subtracted from the SS control sample in this procedure.

Systematic uncertainties are represented through nuisance parameters in the fit, and account for the effects listed in section 8.9, as well as for the energy scale of tag electrons, which is changed by its uncertainty of $\pm1\%$ in the barrel region ($|\eta| < 1.46$) and $\pm2.5\%$ in the endcap regions ($|\eta| > 1.56$), with the difference in the $m_{\text{vis}}$ template considered as an uncertainty in the differential distribution. Similarly, the energy scale of probe electrons and $\tau_h$ are changed by $\pm1.5$ and $\pm3\%$, respectively. The energy scale of leptons have been measured using the method described in ref. [75]. Uncertainties in the normalization of $W$+jets and QCD multijet production are dominated by number of events in the relevant control regions, and each amount to 20%. Finally, an additional 3% uncertainty is associated with the $Z/\gamma^* \rightarrow ee$ normalization because of the need to disentangle possible differences between the $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \tau\tau$ normalizations. Separate fits are used for probes in the barrel and in the endcap regions.

The fitted $m_{\text{vis}}$ distributions in the passing category are shown in figure 14 for the medium and very tight WPs of the against-$e$ discriminant in the barrel region of the ECAL, while the $e \rightarrow \tau_h$ misidentification probabilities are displayed in figure 15. In the barrel region, the measured
Figure 13. Probabilities for quark and gluon jets in $e\mu$+jets events to pass the loose (uppermost), medium (middle), and tight (lowest) WPs of the MVA-based isolation discriminant as a function of $p_T^\text{jet}$ (left) and $\eta^\text{jet}$ (right). The misidentification probabilities in data are compared to expectations from simulation. The vertical bars in the simulated and observed misidentification probabilities include statistical uncertainties from the limited event count in both data and simulated samples, including the background subtraction. The shaded bands contain the systematic uncertainties related to background subtraction and the jet energy scale.
Figure 14. Fitted distributions in $m_{\text{vis}}$ in the passing category for the medium (left) and very tight (right) WPs of the against-$e$ discriminant in the barrel region. The electroweak background includes contributions from W+jets (dominating), diboson, and single top quark events. Vertical bars correspond to the small (not visible) statistical uncertainties in the data points (68% frequentist confidence intervals), while the shaded bands provide the quadratic sum of the statistical and systematic uncertainties after the fit.

Figure 15. Probability for electrons to pass different working points of the against-$e$ discriminant, split into the barrel (left) and endcap (right) regions. For each working point, the $e \to \tau_h$ misidentification probability is defined as the fraction of probes passing that working point relative to the total number of probes. Vertical bars correspond to the statistical and the quadratic sum of the statistical and systematic uncertainties, respectively, for simulated and observed data.
misidentification probabilities in data exceed those in the simulations. The difference between data and simulation increases for the tight and very tight WPs of the discriminant, and a similar trend is observed for the probes in the endcap regions. The observed misidentification probabilities range from \( \approx 5\% \) for the very loose WP to less than 0.1\% for the very tight WP in the barrel region, while in the endcap regions, the probabilities are larger, ranging between 0.1 and 10\%.

### 11.2 Measurement of the \( \mu \rightarrow \tau_h \) probability

The \( \mu \rightarrow \tau_h \) misidentification probability is also measured using a tag-and-probe method, following an approach similar to that used to measure the \( e \rightarrow \tau_h \) misidentification probability discussed in section 11.1. For this, we select \( Z/\gamma^* \rightarrow \mu\mu \) events, as described in section 8.8, and again divide these into two categories, depending on whether the probe passes or fails the specific working point of the against-\( \mu \) discriminant. The number of \( Z/\gamma^* \rightarrow \mu\mu \) signal events in each category is then extracted from a simultaneous maximum likelihood fit to the mass of the tag-and-probe pair, in the range \( 70 < m_{\text{vis}} < 120 \text{GeV} \). Separate fits are performed for probes in five \( |\eta| \) regions of \(<0.4, 0.4–0.8, 0.8–1.2, 1.2–1.7, \text{ and } >1.7 \), corresponding to the geometry of the CMS muon spectrometer.

The normalization and distribution in \( m_{\text{vis}} \) for signal and background processes are estimated as discussed in section 11.1. Systematic uncertainties are also similar, except that those related to electrons are replaced by those appropriate for the muons, such as the energy scale for the probe, which is changed by \( \pm 1.5 \) and \( \pm 3\% \) for the misidentified \( \mu \rightarrow \tau_h \) and the genuine \( \tau_h \) candidates, respectively, with the resulting difference in the \( m_{\text{vis}} \) template taken as an uncertainty in the differential distribution. The uncertainty in the energy scale of the tag muon is negligible compared with the energy scale of the \( \tau_h \) candidates, and is therefore neglected.

Figure 16 shows the mass distribution in the \( \mu\tau_h \) pair after the maximum likelihood fit, for events where the probe muon is reconstructed as a \( \tau_h \) candidate, and passes the loose or tight WPs of the against-\( \mu \) discriminant. The probes in these distributions lie within \( |\eta| < 0.4 \).

The \( \mu \rightarrow \tau_h \) misidentification rates are given for the loose and tight WPs of the against-\( \mu \) discriminant in figure 17. For probes passing the WPs, the measured misidentification rates in data exceed the predictions, with the difference between data and simulation possibly increasing from small to large \( |\eta| \). The observed trend is more significant for probes passing the tight WP. The observed misidentification probabilities for the loose WP are in the range of 0.1–0.5\%, with the highest probability lying in the \( |\eta| \) range between 0.8 and 1.2 which corresponds to transition between barrel and endcap regions of the muon spectrometer. The probabilities for the tight WP range between 0.03 and 0.40\%, with the highest value again falling in the same \( |\eta| \) region.

### 12 Measurement of the \( \tau_h \) energy scale

The correction to the \( \tau_h \) energy scale is defined by the deviation of the average reconstructed \( \tau_h \) energy from the generator-level energy of the visible \( \tau_h \) decay products. The corresponding data-to-simulation correction is obtained from a fit of the distributions of observables sensitive to the energy scale, using samples of \( \tau_h \) and \( \mu\tau_h \) final states in \( Z/\gamma^* \) events. The distributions sensitive to the energy scale are \( m_{\tau_h} \) and the mass of the \( \ell\tau_h \) system, \( m_{\text{vis}} \). These are fitted, separately for the \( h^+, h^+\pi^0 \), and \( h^+h^+h^-h^- \) decays to extract the correction factors between data and simulation.
Figure 16. Fitted distribution in $m_{\text{vis}}$ in the passing category for the loose (left) and tight (right) WPs of the against-$\mu$ discriminant in the region of $|\eta| < 0.4$. The electroweak background includes contributions from W+jets (dominating), diboson, and single top quark events. Vertical bars correspond to the small (not visible) statistical uncertainties in the data points (68% frequentist confidence intervals), while the shaded bands provide the quadratic sum of the statistical and systematic uncertainties after the fit.

Figure 17. Probability for muons to pass the loose (left) and tight (right) WPs of the against-$\mu$ discriminant, as a function of the $|\eta|$ of the probe. For each working point, the $\mu \rightarrow \tau_b$ misidentification probability is defined as the fraction of probes passing that working point relative to the total number of probes. Vertical bars correspond to the statistical and the quadratic sum of the statistical and systematic uncertainties, respectively, for simulated and observed data.
The $e\tau_h$ and $\mu\tau_h$ final states are selected as described in section 8.1, except that the $\tau_h$ candidates are required to pass the very tight WP of the MVA-based $\tau_h$ isolation discriminant to further reduce backgrounds from jets misidentified as $\tau_h$ candidates. Moreover, the requirement on $m_T$ is tightened to be less than 30 GeV, and the requirement on $D_{\tau}$ is removed. Finally, the $\tau_h$ candidate must satisfy the tight and loose, or very loose and tight WPs of the against-$e$ and against-$\mu$ discriminants in the respective $e\tau_h$ or $\mu\tau_h$ final states. Templates for events in which the reconstructed $\tau_h$ is matched to some generated $\tau_h$ are obtained by changing the reconstructed $\tau_h$ energy between $-6\%$ and $+6\%$ in steps of 0.1%, with the $m_{\text{vis}}$ and $m_{\tau_h}$ recomputed at each step. The maximal energy shifts of $\pm6\%$ are selected to be sufficiently away from the nominal value in the simulation such that the true value can be obtained between them. While $m_{\tau_h}$ displays higher sensitivity to the energy scale for the $h^{+}\pi^{0}$ and $h^{+}h^{-}h^{+}$ decay modes, it cannot be used in the $h^{+}$ decay mode, where only $m_{\text{vis}}$ is used. The backgrounds are modelled in the same way as described in section 9.1, and the templates for processes in which there is no match between the reconstructed and generated $\tau_h$ candidates are not changed as a function of the $\tau_h$ energy scale.

For illustration, the $m_{\tau_h}$ templates corresponding to no shifts, and to shifts in $\tau_h$ energy scale of $-6$ and $+6\%$ are shown in figure 18 for events selected in $h^{+}\pi^{0}$ decay mode. The data are compared to predictions for these three energy scales.

A likelihood ratio method is used to extract the $\tau_h$ energy scale for each decay mode. In addition to those listed in section 8.9, the following sources of systematic uncertainties are considered: uncertainties in the identification of $\tau_h$ candidates, determined in section 9, are split into those that are uncorrelated ($\approx2\%$) and correlated ($\approx4.5\%$) between the $e\tau_h$ and $\mu\tau_h$ final states. The rates for electrons, muons, and jets misidentified as $\tau_h$ candidates have uncertainties of 12, 25, and 20%, respectively. Moreover, uncertainties in the energy scale of electrons (1% in the barrel and 2.5% in the endcaps) and muons (5%) identified as $\tau_h$ are taken into account in their differential distributions. The results obtained from fits to $m_{\text{vis}}$ and $m_{\tau_h}$ distributions for each decay mode in the $e\tau_h$ and $\mu\tau_h$ final states are found to be compatible with each other, and their combination is given in table 7. The measurement is limited by systematic rather than statistical uncertainties.

Table 7. The data-to-simulation correction for the $\tau_h$ energy scale from the combination of measurements performed in the $e\tau_h$ and $\mu\tau_h$ final states separately using $m_{\tau_h}$ and $m_{\text{vis}}$ distributions. The correction is relative to the reconstructed energy from simulation, expressed in %.

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>$m_{\tau_h}$</th>
<th>$m_{\text{vis}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h^{+}$</td>
<td>$-0.5\pm0.5$</td>
<td></td>
</tr>
<tr>
<td>$h^{+}\pi^{0}$</td>
<td>$0.9\pm0.3$</td>
<td>$1.1\pm0.3$</td>
</tr>
<tr>
<td>$h^{+}h^{-}h^{+}$</td>
<td>$0.6\pm0.2$</td>
<td>$0.6\pm0.3$</td>
</tr>
</tbody>
</table>

Additional studies performed using the $\mu\tau_h$ final state are carried out to assess the stability of the measurement. To gauge the impact of fluctuations caused by the limited number of MC events relative to the data, the simulated events used to model $Z/\gamma^{*}$ decays are split into four samples of equal size, and the measurement is performed using each of these four subsamples. The resulting fluctuations in the measured $\tau_h$ energy scale are up to 1%. Similarly, the effect of the contamination from backgrounds that arise from misidentification of the $\tau_h$ is checked by changing the selection criteria, and found to be 0.5%. The choice of the binning is investigated by changing the number
Figure 18. The distributions in $m_{\tau_h}$ for $\mu \tau_h$ events in the $h^0 \pi^0$ decay channel. The data are compared to predictions with different shifts applied to the $\tau_h$ energy scale: 0% (upper), −6% (lower left), and +6% (lower right). The electroweak background includes contributions from $W$+jets (dominating), diboson, and single top quark events. Vertical bars (smaller than the symbol size) correspond to the statistical uncertainty in the data points (68% frequentist confidence intervals), while the shaded bands provide the expected systematic uncertainties.

of bins up and down by a factor 2. The results are compatible to within 1%. Finally, the effect of the range of the fit is evaluated for the $m_{vis}$ template by increasing it by 10 GeV in either direction, resulting in changes compatible within 0.5% of the original measurement. Although these checks do not guarantee that similar levels of fluctuation exist in the original measurement (especially, the assessment of the limited number of MC events), an additional uncertainty of 1.0% is added in quadrature to the uncertainty detailed in table 7, to reflect our limited knowledge of the true fluctuations. This results in a total uncertainty of <1.2%.

13 Performance of $\tau_h$ identification in the high-level trigger

The $\tau_h$ reconstruction and identification algorithm described in section 7 for the HLT was used to define a set of triggers for 2016 data taking. These triggers cover all final states of interest, namely,
τ lepton pair production in ττ, τµτ, and τhτh decays, τh associated with pTmiss (τhpTmiss), and single τh with large pT.

There are two types of HLT decision trees that use τh candidates and which are aimed at two different classes of final states, those that comprise other than τh candidates in the event, e.g., eτh, µτh, τhpmiss, and those that include only τh candidates, e.g., τhτh. The first type of trigger is based on L1 seeds that require the presence of an electron, a muon, or large pTmiss, possibly together with a τh candidate. These triggers also have their corresponding selections in e, µ, or pTmiss in the HLT, thereby greatly reducing the event rates processed at later stages. This allows reconstruction of τh candidates directly through the resource-intensive L3 step, wherein the PF sequence underpinning τh reconstruction is run using the full-detector acceptance. In the second type of trigger, only τh candidates are required to be identified at L1, without additional lepton or pTmiss selections. At HLT, since the L3 step would be too time consuming to run at the L1 output rates, the L2 and L2.5 filtering steps are executed first. The efficiency of the L2 and L2.5 filter is >95% per τh candidate. In addition, this class of triggers has HLT τh reconstruction used only in regions of the detector centered on the direction of the L1 τh candidates, further reducing thereby the processing time.

The triggers for τ pair production are aimed mainly to select efficiently the SM H → ττ decays that require respective pT thresholds of 20–25 and 30–35 GeV for τe or τµ and τh final states. In addition, trigger rates at an instantaneous luminosity of L = 1.4 × 10^34 cm^{-2} s^{-1} and PU close to 40 interactions per bunch crossing, typical for pp collisions in late 2016, were required not to exceed rates of about 10–15 and 50–65 Hz for the eτh or µτh and τhτh triggers, respectively.

The µτh trigger is constructed as follows. First, we require the presence of a muon candidate with pT > 18 GeV at L1. Then, an isolated muon, seeded by the L1 candidate, with pT > 19 GeV is selected at the HLT. Subsequently, an unseeded L3 τh candidate is selected with pT > 20 GeV that passes the loose charged-particle isolation WP. The isolation is relaxed linearly by 10%/GeV for pTmiss > 50 GeV. Finally, the L3 τh candidate must be separated from the muon by ΔR > 0.3. At $L = 1.4 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$, the rate for the µτh trigger corresponds to ~20 Hz.

To adapt to different conditions in instantaneous luminosity delivered by the LHC in 2016, ranging from ~3 × 10^33 cm^{-2} s^{-1} to 1.4 × 10^{34} cm^{-2} s^{-1}, and to provide highest efficiency possible within the limited rate budget, two variants of eτh triggers were developed. The first one is similar to the µτh trigger in that an isolated electromagnetic (e or γ) object with pT > 22 GeV is required at L1, and is used to initiate the reconstruction of an isolated electron at the HLT that is required to have pT > 24 GeV. A seedless L3 τh candidate, not overlapping with the electron, is required to have pT > 20 GeV, and to pass the loose charged-particle isolation WP (linearly relaxed by 10%/GeV for pTmiss > 50 GeV). This trigger covered instantaneous luminosities of up to 9 × 10^{33} cm^{-2} s^{-1}.

The second, a more stringent version of the eτh trigger, adds the requirement of an L1 τh candidate to accompany the L1 electromagnetic object. First, the pT threshold on the L1 τh was set to 20 GeV, and, as the instantaneous luminosity increased, was raised to 26 GeV, and eventually the L1 isolation condition was also applied. In the latter configuration at the HLT, the pT threshold for the L3 τh candidate was adjusted to 30 GeV. In the utilized ranges of instantaneous luminosity for which the eτh triggers were designed, the trigger rates remained below 15 Hz.

The τhτh triggers require a pair of isolated L1 τh candidates, with pT above a threshold in the range of 28–36 GeV. The threshold is dynamically adjusted to maintain a constant rate of events passing L1, independent of the instantaneous luminosity. Even after satisfying the L1 requirements,
the event rate is still too high to run the L3 $\tau_h$ reconstruction. The L3 reconstruction is therefore used only if at least two $\tau_h$ candidates pass the L2 and L2.5 stages, as discussed in section 7. At L3, the candidates have to have $p_T > 35$ GeV, and pass the medium WP of the charged isolation (the charged isolation was replaced by the combined isolation at $L > 1.3 \times 10^{34}$ cm$^{-2}$ s$^{-1}$). The isolation is linearly relaxed by 6%/GeV for $p_T^{\tau_h} > 73$ GeV. Two such candidates must be present in the event, and must be separated by $\Delta R > 0.5$. At $L = 1.4 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, the rate of $\tau_h \tau_h$ triggers was below 60 Hz.

The benchmark process that guided the design of the $\tau_h p_T^{\text{miss}}$ trigger is the decay of a charged resonance $X^* \rightarrow \tau \nu$, e.g., for $X^* = H^*$ or $W'$, with a mass $m_X > 200$ GeV. At L1, this trigger requires $p_T^{\text{miss}}$ in excess of 80–100 GeV, again with the threshold dynamically adjusted as a function of instantaneous luminosity to keep the rate of events passing L1 constant. At the HLT, the selected events must further satisfy the condition of $p_T^{\text{miss}} > 90$ GeV. After this, the L3 $\tau_h$ reconstruction step is executed, and events are finally saved when an L3 $\tau_h$ candidate with $p_T > 50$ GeV, passing the loose WP of charged isolation (relaxed by 6%/GeV for $p_T^{\tau_h} > 100$ GeV) is found, with its leading charged hadron satisfying $p_T > 30$ GeV. At $L = 1.4 \times 10^{34}$ cm$^{-2}$ s$^{-1}$, the rate for the $\tau_h p_T^{\text{miss}}$ trigger is about 20 Hz.

Finally, a high-$p_T$ single-$\tau_h$ trigger was developed for searches for high-mass resonances decaying into at least one $\tau$ lepton, for example $W^*, H^*$, or the heavy A or H boson in MSSM. This trigger was designed to cover portions of the phase space not covered by the more usual cross-triggers ($\tau_h \tau_h$, $\tau_h p_T^{\text{miss}}$, $e\tau_h$, and $\mu \tau_h$), e.g., $H^*$ events with an energetic $\tau_h$ but small $p_T^{\text{miss}}$. The trigger that fulfilled those conditions required an isolated L1 $\tau_h$ candidate with $p_T > 120$ GeV. The $\tau_h$ reconstruction at the HLT consists of steps taken in L2, L2.5, and L3. The L3 requires one $\tau_h$ candidate with $p_T > 140$ GeV, and with a leading charged hadron with $p_T > 50$ GeV. The L3 $\tau_h$ candidate has to pass also the tight WP of charged isolation, which is linearly relaxed by 2%/GeV for $p_T^{\tau_h} > 275$ GeV, and is discarded for $p_T^{\tau_h} > 500$ GeV. Rates of about 30 Hz were allocated to this trigger.

The basic features of triggers with $\tau_h$ candidates used to record pp collisions in 2016 are summarized in table 8. The efficiencies of the $\tau_h$ part of the triggers listed in table 8 are measured via the tag-and-probe technique as a function of the offline-reconstructed $p_T^{\tau_h}$, using data enriched in $\tau_h$ leptons from $Z/\gamma^* \rightarrow \tau \tau \rightarrow \mu \tau_h$ decays. To single out this sample, the selections for the $\mu \tau_h$ final state described in section 8.1, together with the requirement of $m_{\tau_h} < 30$ GeV and the additional condition of $40 < m_{\text{vis}} < 80$ GeV, are applied to data previously collected through single-muon triggers. Furthermore, to provide an efficiency measurement that is specific to the selections used in $H \rightarrow \tau \tau$ analyses, the $\tau_h$ candidates must pass the tight WP of the MVA-based isolation discriminant. The residual contamination from other objects misidentified as $\tau_h$ is subtracted statistically using SS events passing the same selections. The purity of the final sample exceeds 95%.

To provide an unbiased measurement of the efficiency of the single-$\tau_h$ part of the $\tau_h \tau_h$ and $\tau_h p_T^{\text{miss}}$ triggers, special $\mu \tau_h$ triggers were put in place. The special triggers have one part that is required to match the nominal single-muon trigger used to select events; the other part is required to pass the $\tau_h$ trigger identification for the trigger of interest.

The passing $\tau_h$ probes correspond to those that pass the $\tau_h$ part of the special trigger, i.e., the trigger is satisfied and its $\tau_h$ part geometrically matches ($\Delta R < 0.5$) the selected offline $\tau_h$. The efficiency of the $\tau_h$ part of the $\mu \tau_h$ and $\tau_h \tau_h$ triggers, measured using collision data relative to the DY simulation, is shown in figure 19. For the $\tau_h \tau_h$ trigger, we use only the portion of the 2016 data that contains the trigger employing the combined isolation. In both cases, simulation agrees well with the data. Data-to-simulation agreement is similar for the other triggers discussed in this section.
Table 8. Triggers with $\tau_h$ candidates used to record pp collisions in 2016: the final state (Channel), HLT $p_T$ thresholds and $\tau_h$ isolation working point, L1 $p_T$ thresholds, peak instantaneous luminosity ($L_{\text{peak}}$) in the period of operation as main trigger, and integrated luminosity ($\int L$) collected with the trigger. The $\tau_h\mu$ and $\tau_h\mu^{\text{miss}}$ triggers are seeded by sets of L1 triggers with thresholds dynamically adjusted as a function of the instantaneous luminosity to maintain a constant L1 rate, given by the ranges in $p_T$. The trigger $p_T$ thresholds and isolation criteria were successively tightened over the data-taking period to keep the rate of events passing HLT approximately constant with increasing instantaneous luminosity.

<table>
<thead>
<tr>
<th>Channel</th>
<th>HLT object and WP</th>
<th>L1 object</th>
<th>$L_{\text{peak}}$ (cm$^{-2}$ s$^{-1}$)</th>
<th>$\int L$ (fb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu\tau_h$</td>
<td>$p_T^{\mu} &gt; 19$ GeV, $p_T^{\tau} &gt; 20$ GeV, loose iso</td>
<td>$p_T^{\mu} &gt; 18$ GeV</td>
<td>$1.5 \times 10^{34}$</td>
<td>35.9</td>
</tr>
<tr>
<td>$e\tau_h$</td>
<td>$p_T^e &gt; 24$ GeV, $p_T^{\tau} &gt; 20$ GeV, loose iso</td>
<td>$p_T^{\tau} &gt; 22$ GeV</td>
<td>$0.9 \times 10^{34}$</td>
<td>7.5</td>
</tr>
<tr>
<td>$e\tau_h$</td>
<td>$p_T^e &gt; 24$ GeV, $p_T^{\tau} &gt; 20$ GeV, loose iso</td>
<td>$p_T^{\tau} &gt; 22$ GeV, $p_T^{\mu} &gt; 20$ GeV</td>
<td>$1.3 \times 10^{34}$</td>
<td>10.2</td>
</tr>
<tr>
<td>$\tau_h\tau_h$</td>
<td>$2 \times p_T^{\tau} &gt; 35$ GeV, medium iso</td>
<td>$2 \times \text{iso } p_T^{\tau} &gt; 28$–36 GeV</td>
<td>$1.3 \times 10^{34}$</td>
<td>27.3</td>
</tr>
<tr>
<td>$\tau_h\tau_h$</td>
<td>$2 \times p_T^{\tau} &gt; 35$ GeV, medium comb. iso</td>
<td>$2 \times \text{iso } p_T^{\tau} &gt; 28$–36 GeV</td>
<td>$1.5 \times 10^{34}$</td>
<td>8.6</td>
</tr>
<tr>
<td>$\tau_h\mu^{\text{miss}}$</td>
<td>$p_T^{\mu^{\text{miss}}} &gt; 90$ GeV, $p_T^{\tau} &gt; 50$ GeV, $p_T^{\mu} &gt; 30$ GeV, loose iso</td>
<td>$p_T^{\mu^{\text{miss}}} &gt; 80$–100 GeV</td>
<td>$1.5 \times 10^{34}$</td>
<td>35.9</td>
</tr>
<tr>
<td>$\tau_h$</td>
<td>$p_T^{\tau} &gt; 140$ GeV, $p_T^{\mu} &gt; 50$ GeV, tight iso</td>
<td>$p_T^{\tau} &gt; 120$ GeV</td>
<td>$1.5 \times 10^{34}$</td>
<td>33.1</td>
</tr>
</tbody>
</table>

Figure 19. Single-$\tau_h$ efficiency of the $\mu\tau_h$ (left) and $\tau_h\tau_h$ (right) triggers. The efficiency is computed per single $\tau_h$, using the tag-and-probe method as a function of the offline-reconstructed $p_T^{\tau_h}$. Observed data are compared to simulated $Z/\gamma^* \rightarrow \tau\tau$ events selected through the same procedure. Vertical bars correspond to the statistical uncertainties. The plot on the right has data points fitted using a cumulative (integral) distribution of the Crystal Ball function [76].

Figure 19 shows that the nominal $p_T$ threshold of the $\tau_h$ triggers corresponds to an efficiency of 50%, as expected for trigger and offline objects with the same energy scale. The slow turn-on originates from two effects: in the $p_T$ range above about twice the trigger threshold, it is caused by the relaxed isolation selection applied at HL, but not in the offline selection; in the range just above the trigger threshold, it is caused by an asymmetric energy response of the HLT $\tau_h$ candidate relative to its offline counterpart. The asymmetry is due to a more inclusive selection of constituents.
of the $\tau_h$ candidate at the HLT than offline. The second effect is clearly visible in the $\mu\tau_h$ trigger with unseeded L3 $\tau_h$ reconstruction, while for the $\tau_\eta\tau_h$ trigger it is smeared out by the resolution of the L1- and L2-$\tau_\eta$ candidates (relative to offline), which is much worse than the resolution of the L3 candidates.

14 Summary

The “hadron-plus-strips” algorithm developed at the CMS experiment to reconstruct and identify $\tau \rightarrow \text{hadrons} + \nu_\tau$ decays in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV, as presented in ref. [24], has been improved. The changes include a dynamic strip reconstruction, the reconstruction of highly boosted $\tau$ lepton pairs, and the introduction of additional variables in the multivariate-analysis discriminants used to reject jets and electrons. The isolation discriminants have also been optimized to cope with the large pileup of events in $\sqrt{s} = 13$ TeV proton-proton runs.

The performance of the improved algorithm has been measured using 35.9 fb$^{-1}$ of data recorded during 2016 at $\sqrt{s} = 13$ TeV. The $\tau_h$ identification efficiency in data at low, intermediate, and high transverse momenta, as well as for highly Lorentz-boosted $\tau$ lepton pairs, is similar to that expected from Monte Carlo simulation, while differences of 10–20% are found between data and simulation for the jet $\rightarrow \tau_h$ misidentification probability. The $e \rightarrow \tau_h$ and $\mu \rightarrow \tau_h$ misidentification probabilities are smaller than those of the previous version of the algorithm under the same running conditions, while maintaining a high efficiency for the selection of genuine $\tau_h$ candidates. The corresponding data-to-simulation scale factors have also been determined. The energy scale of $\tau_h$ candidates is measured, and its response relative to Monte Carlo simulation is found to be close to unity. Finally, a specialized $\tau_h$ reconstruction and identification algorithm has been used in the high-level trigger, and its performance has been presented.

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Education, Science and Research and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Secretariat for Higher Education, Science, Technology and Innovation, Ecuador; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture,
contracts Harmonia 2014/14/M/ST2/00428, Opus 2014/13/B/ST2/02543, 2014/15/B/ST2/03998, and 2015/19/B/ST2/02861, Sonata-bis 2012/07/E/ST2/01406; the National Priorities Research Program by Qatar National Research Fund; the Programa de Excelencia María de Maeztu, and the Programa Severo Ochoa del Principado de Asturias; the Thalis and Aristeia programmes cofinanced by EU-ESF, and the Greek NSRF; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University, and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Weston Havens Foundation (U.S.A.).

References


[8] CMS collaboration, Searches for a heavy scalar boson \( H \) decaying to a pair of 125 GeV Higgs bosons \( hh \) or for a heavy pseudoscalar boson \( A \) decaying to \( Zh \), in the final states with \( h \rightarrow \tau\tau \), Phys. Lett. B 755 (2016) 217 [arXiv:1510.01181].

[9] CMS collaboration, Search for lepton flavour violating decays of the Higgs boson to \( e\tau \) and \( e\mu \) in proton-proton collisions at \( \sqrt{s} = 8 \) TeV, Phys. Lett. B 763 (2016) 472 [arXiv:1607.03561].


[61] C.J. Clopper and E.S. Pearson, The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika 26 (1934) 404.


The CMS collaboration

Yerevan Physics Institute, Yerevan, Armenia
A.M. Sirunyan, A. Tumasyan

Institut für Hochenergiephysik, Wien, Austria

Institute for Nuclear Problems, Minsk, Belarus
V. Chekhovsky, V. Mossolov, J. Suarez Gonzalez

Universiteit Antwerpen, Antwerpen, Belgium

Vrije Universiteit Brussel, Brussel, Belgium

Université Libre de Bruxelles, Bruxelles, Belgium

Ghent University, Ghent, Belgium

Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
Universidade Estadual Paulista $^{a}$, Universidade Federal do ABC $^{b}$, São Paulo, Brazil
S. Ahuja$^{a}$, C.A. Bernardes$^{a}$, L. Calligaris$^{a}$, T.R. Fernandez Perez Tomei$^{a}$, E.M. Gregores$^{b}$, P.G. Mercadante$^{b}$, S.F. Novaes$^{a}$, Sandra S. Padula$^{a}$

Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria
A. Aleksandrov, R. Hadjiiska, P. Iaydjiev, A. Marinov, M. Misheva, M. Rodozov, M. Shopova, G. Sultanov

University of Sofia, Sofia, Bulgaria
A. Dimitrov, L. Litov, B. Pavlov, P. Petkov

Beihang University, Beijing, China
W. Fang$^{5}$, X. Gao$^{5}$, L. Yuan

Institute of High Energy Physics, Beijing, China

State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
Y. Ban, G. Chen, A. Levin, J. Li, L. Li, Q. Li, Y. Mao, S.J. Qian, D. Wang, Z. Xu

Tsinghua University, Beijing, China
Y. Wang

Universidad de Los Andes, Bogota, Colombia
C. Avila, A. Cabrera, C.A. Carrillo Montoya, L.F. Chaparro Sierra, C. Florez, C.F. González Hernández, M.A. Segura Delgado

University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia
B. Courbon, N. Godinovic, D. Lelas, I. Puljak, T. Sculac

University of Split, Faculty of Science, Split, Croatia
Z. Antunovic, M. Kovac

Institute Rudjer Boskovic, Zagreb, Croatia
V. Brigljevic, D. Ferencek, K. Kadija, B. Mesic, A. Starodumov$^{7}$, T. Susa

University of Cyprus, Nicosia, Cyprus

Charles University, Prague, Czech Republic
M. Finger$^{8}$, M. Finger Jr.$^{8}$

Escuela Politecnica Nacional, Quito, Ecuador
E. Ayala
Universidad San Francisco de Quito, Quito, Ecuador
E. Carrera Jarrin

Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
A. Mahrous9, A. Mohamed10, E. Salama11,12

National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
S. Bhowmik, A. Carvalho Antunes De Oliveira, R.K. Dewanjee, K. Ehatat, M. Kadastik, M. Raidal, C. Veelken

Department of Physics, University of Helsinki, Helsinki, Finland
P. Eerola, H. Kirschenmann, J. Pekkanen, M. Voutilainen

Helsinki Institute of Physics, Helsinki, Finland

Lappeenranta University of Technology, Lappeenranta, Finland
T. Tuuva

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

Laboratoire Leprince-Ringuet, Ecole polytechnique, CNRS/IN2P3, Université Paris-Saclay, Palaiseau, France

Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France

Centre de Calcul de l’Institut National de Physique Nucleaire et de Physique des Particules, CNRS/IN2P3, Villeurbanne, France
S. Gadrat

Université de Lyon, Université Claude Bernard Lyon 1, CNRS-IN2P3, Institut de Physique Nucléaire de Lyon, Villeurbanne, France
Georgian Technical University, Tbilisi, Georgia
T. Toriashvili

Tbilisi State University, Tbilisi, Georgia
Z. Tsamalaidze

RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany

RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany

RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany

Deutsches Elektronen-Synchrotron, Hamburg, Germany

University of Hamburg, Hamburg, Germany

Karlsruher Institut fuer Technologie, Karlsruhe, Germany
Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece

National and Kapodistrian University of Athens, Athens, Greece

National Technical University of Athens, Athens, Greece
K. Kousouris, I. Papakrivopoulos, G. Tsiplitis

University of Ioánnina, Ioánnina, Greece

MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary

Wigner Research Centre for Physics, Budapest, Hungary

Institute of Nuclear Research ATOMKI, Debrecen, Hungary
N. Beni, S. Czellar, J. Karancsi, A. Makovec, J. Molnar, Z. Szillasi

Institute of Physics, University of Debrecen, Debrecen, Hungary
P. Raics, Z.L. Trocsanyi, B. Ujvari

Indian Institute of Science (IISc), Bangalore, India
S. Choudhury, J.R. Komaragiri, P.C. Tiwari

National Institute of Science Education and Research, HBNI, Bhubaneswar, India

Panjab University, Chandigarh, India

University of Delhi, Delhi, India
A. Bhardwaj, B.C. Choudhary, R.B. Garg, M. Gola, S. Keshri, Ashok Kumar, S. Malhotra, M. Naimuddin, P. Priyanka, K. Ranjan, Aashaq Shah, R. Sharma

Saha Institute of Nuclear Physics, HBNI, Kolkata, India

Indian Institute of Technology Madras, Madras, India
P.K. Behera
Bhabha Atomic Research Centre, Mumbai, India
R. Chudasama, D. Dutta, V. Jha, V. Kumar, P.K. Netrakanti, L.M. Pant, P. Shukla

Tata Institute of Fundamental Research-A, Mumbai, India
T. Aziz, M.A. Bhat, S. Dugad, G.B. Mohanty, N. Sur, B. Sutar, RavindraKumar Verma

Tata Institute of Fundamental Research-B, Mumbai, India
S. Banerjee, S. Bhattacharya, S. Chatterjee, P. Das, M. Guchait, S. Jain, S. Karmakar, S. Kumar, M. Maity

Indian Institute of Science Education and Research (IISER), Pune, India
S. Chauhan, S. Dube, V. Hegde, A. Kapoor, K. Kothekar, S. Pandey, A. Rane, S. Sharma

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
S. Chenarani, E. Eskandari Tadavani, S.M. Etesami, M. Khakzad, M. Mohammadi Najafabadi, M. Naseri, F. Rezaei Hosseinabadi, M. Zeinali

University College Dublin, Dublin, Ireland
M. Felcini, M. Grunewald

INFN Sezione di Bari a, Università di Bari b, Politecnico di Bari c, Bari, Italy

INFN Sezione di Bologna a, Università di Bologna b, Bologna, Italy

INFN Sezione di Catania a, Università di Catania b, Catania, Italy
S. Albergo, A. Di Mattia, R. Potenza, A. Tricomi, C. Tuve

INFN Sezione di Firenze a, Università di Firenze b, Firenze, Italy

INFN Laboratori Nazionali di Frascati, Frascati, Italy
L. Benussi, S. Bianco, F. Fabbri, D. Piccolo

INFN Sezione di Genova a, Università di Genova b, Genova, Italy
F. Ferro, F. Ravera, E. Robutti, S. Tosi
INFN Sezione di Milano-Bicocca a, Università di Milano-Bicocca b, Milano, Italy
A. Benaglia a, A. Beschb, L. Brianza a,b, F. Brivio a,b, V. Cirio a,b, S. Di Guida a,d,17, M.E. Dinardo a,b, S. Fiorendi a,b, S. Gennai a, A. Ghezzi a,b, P. Govoni a,b, M. Malberti a,b, S. Malvezzi a, A. Massironi a,b, D. Menasse a, F. Monti, L. Moroni a, M. Paganoni a,b, D. Pedrini a, S. Ragazzi a,b, T. Tabarelli de Fatis a,b, D. Zuo a,b

INFN Sezione di Napoli a, Università di Napoli ‘Federico II’ b, Napoli, Italy, Università della Basilicata c, Potenza, Italy, Università G. Marconi a, Roma, Italy
S. Buontempo a, N. Cavallo a,c, A. Di Crescenzo a,b, F. Fabozzi a,c, F. Fienga a, G. Galati a, A.O.M. Iorio a,b, W.A. Khan a, L. Lista a, S. Meola a,d,17, P. Paolucci a,17, C. Sciaccia a,b, E. Voevodina a,b

INFN Sezione di Padova a, Università di Padova b, Padova, Italy, Università di Trento c, Trento, Italy
P. Azzi a, N. Bacchetta a, A. Boletti a,b, A. Bragagnolo, R. Carlin a,b, P. Checchia a, M. Dall’Osso a,b, P. De Castro Manzano a, T. Dorigo a, U. Dosselli a, F. Gasparini a,b, U. Gasparini a,b, A. Gozzelino a, S.Y. Hoh, S. Lacaprara a, P. Lujan, M. Margon a,b, A.T. Meneguzzo a,b, M. Passaeo a, J. Pazzini a,b, N. Pozzobon a,b, P. Ronchese a,b, R. Rossini a,b, F. Simonetto a,b, A. Tiko, E. Torass a, M. Zanetti a,b, P. Zotto a,b

INFN Sezione di Pavia a, Università di Pavia b, Pavia, Italy
A. Braghieri a, A. Magnani a, P. Montagna a,b, S.P. Ratti a,b, V. Re a, M. Ressegotti a,b, C. Riccardi a,b, P. Salvini a, I. Vai a,b, P. Vitulo a,b

INFN Sezione di Perugia a, Università di Perugia b, Perugia, Italy
M. Biasini a,b, G.M. Bilei a, C. Cecchi a,b, D. Ciangottini a,b, L. Fanò a,b, P. Lariccia a,b, R. Leonardi a,b, E. Manoni a, G. Mantovani a,b, V. Mariani a,b, M. Menichelli a, A. Rossi a,b, A. Santocchia a,b, D. Spiga a

INFN Sezione di Pisa a, Università di Pisa b, Scuola Normale Superiore di Pisa c, Pisa, Italy
K. Androsov a, P. Azzurri a, G. Bagliesi a, L. Bianchini a, T. Boccali a, L. Borrello, R. Castaldi a, M.A. Ciocci a,b, R. Dell’Orso a, G. Fedi a, F. Fiori a,c, L. Giannini a,c, A. Giassi a, M.T. Gripp a, F. Ligabue a,c, E. Manca a,c, G. Mandorli a,c, A. Messineo a,b, F. Palla a, A. Rizzi a,b, P. Spagnolo a, R. Tenchini a, G. Tonelli a,b, A. Venturi a, P.G. Verdini a

INFN Sezione di Roma a, Sapienza Università di Roma b, Rome, Italy
L. Barone a,b, F. Cavallari a, M. Cipriani a,b, D. Del Re a,b, E. Di Marco a,b, M. Diemoz a, S. Gelli a,b, E. Longo a,b, B. Marzocchi a,b, P. Meridiana a, G. Organtini a,b, F. Pandolfi a, R. Paramatti a,b, F. Preiat a,b, S. Rahatlou a,b, C. Revelli a, F. Santanastasio a,b

INFN Sezione di Torino a, Università di Torino b, Torino, Italy, Università del Piemonte Orientale c, Novara, Italy
N. Amapane a,b, R. Arcidiacono a,c, S. Argiro a,b, M. Arneodo a,c, N. Bartosik a, R. Bellan a,b, C. Biino a, N. Cartiglia a, F. Cenna a,b, S. Cometti a, M. Costa a,b, R. Covarelli a,b, N. Demaria a, B. Kiani a,b, C. Mariotti a, S. Maselli a, E. Migliore a,b, V. Monaco a,b, E. Monteti a,b, M. Monteno a, M.M. Obertino a,b, L. Pacher a,b, N. Pastrone a, M. Pelliccioni a, G.L. Pinna Angioni a,b
A. Romero\textsuperscript{a, b}, M. Ruspa\textsuperscript{a, c}, R. Sacchi\textsuperscript{a, b}, K. Shchelina\textsuperscript{a, b}, V. Sola\textsuperscript{a}, A. Solano\textsuperscript{a, b}, D. Soldi\textsuperscript{a, b}, A. Staiano\textsuperscript{a}

INFN Sezione di Trieste \textsuperscript{a}, Università di Trieste \textsuperscript{b}, Trieste, Italy
S. Belforte\textsuperscript{a}, V. Candelise\textsuperscript{a, b}, M. Casarsa\textsuperscript{a}, F. Cossutti\textsuperscript{a}, A. Da Rold\textsuperscript{a, b}, G. Della Ricca\textsuperscript{a, b}, F. Vazzoler\textsuperscript{a, b}, A. Zanetti\textsuperscript{a}

Kyungpook National University, Daegu, Korea

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Korea
H. Kim, D.H. Moon, G. Oh

Hanyang University, Seoul, Korea
J. Goh\textsuperscript{31}, T.J. Kim

Korea University, Seoul, Korea

Sejong University, Seoul, Korea
H.S. Kim

Seoul National University, Seoul, Korea

University of Seoul, Seoul, Korea

Sungkyunkwan University, Suwon, Korea
Y. Choi, C. Hwang, J. Lee, I. Yu

Vilnius University, Vilnius, Lithuania
V. Dudenas, A. Juodagalvis, J. Vaitkus

National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia

Universidad de Sonora (UNISON), Hermosillo, Mexico
J.F. Benitez, A. Castaneda Hernandez, J.A. Murillo Quijada

Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico
Moscow Institute of Physics and Technology, Moscow, Russia
T. Aushev

National Research Nuclear University ‘Moscow Engineering Physics Institute’ (MEPhI), Moscow, Russia
M. Chadeeva, P. Parygin, D. Philippov, S. Polikarpov, E. Popova, V. Rusinov

P.N. Lebedev Physical Institute, Moscow, Russia
V. Andreev, M. Azarkin, I. Dremin, M. Kirakosyan, S.V. Rusakov, A. Terkulov

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia

Novosibirsk State University (NSU), Novosibirsk, Russia
A. Barnyakov, V. Blinov, T. Dimova, L. Kardapoltsev, Y. Skovpen

Institute for High Energy Physics of National Research Centre ‘Kurchatov Institute’, Protvino, Russia

National Research Tomsk Polytechnic University, Tomsk, Russia
A. Babaev, S. Baidali, V. Okhotnikov

University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
P. Adzic, P. Cirkovic, D. Devetak, M. Dordevic, J. Milosevic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

Universidad Autónoma de Madrid, Madrid, Spain
C. Albajar, J.F. de Trocóniz

Universidad de Oviedo, Oviedo, Spain

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain

University of Ruhuna, Department of Physics, Matara, Sri Lanka

N. Wickramage

CERN, European Organization for Nuclear Research, Geneva, Switzerland


Paul Scherrer Institut, Villigen, Switzerland


ETH Zurich - Institute for Particle Physics and Astrophysics (IPA), Zurich, Switzerland


Universität Zürich, Zurich, Switzerland


National Central University, Chung-Li, Taiwan


National Taiwan University (NTU), Taipei, Taiwan

Chulalongkorn University, Faculty of Science, Department of Physics, Bangkok, Thailand
B. Asavapibhop, N. Srimanobhas, N. Suwonjandee

Çukurova University, Physics Department, Science and Art Faculty, Adana, Turkey

Middle East Technical University, Physics Department, Ankara, Turkey
B. Isildak, G. Karapinar, M. Yalvac, M. Zeyrek

Bogazici University, Istanbul, Turkey
I.O. Atakisi, E. Gülmez, M. Kaya, O. Kaya, S. Ozkorucuklu, S. Tekten, E.A. Yetkin

Istanbul Technical University, Istanbul, Turkey
M.N. Agaras, A. Cakir, K. Cankocak, Y. Komurcu, S. Sen

Institute for Scintillation Materials of National Academy of Science of Ukraine, Kharkov, Ukraine
B. Grynyov

National Scientific Center, Kharkov Institute of Physics and Technology, Kharkov, Ukraine
L. Levcchuk

University of Bristol, Bristol, United Kingdom

Rutherford Appleton Laboratory, Didcot, United Kingdom

Imperial College, London, United Kingdom

Brunel University, Uxbridge, United Kingdom
J.E. Cole, P.R. Hobson, A. Khan, P. Kyberd, C.K. Mackay, A. Morton, I.D. Reid, L. Teodorescu, S. Zahid

Baylor University, Waco, U.S.A.
K. Call, J. Dittmann, K. Hatakeyama, H. Liu, C. Madrid, B. Mcmaster, N. Pastika, C. Smith
Catholic University of America, Washington DC, U.S.A.
R. Bartek, A. Dominguez

The University of Alabama, Tuscaloosa, U.S.A.
A. Buccilli, S.I. Cooper, C. Henderson, P. Rumerio, C. West

Boston University, Boston, U.S.A.
D. Arcaro, T. Bose, D. Gastler, D. Rankin, C. Richardson, J. Rohlf, L. Sulak, D. Zou

Brown University, Providence, U.S.A.

University of California, Davis, Davis, U.S.A.

University of California, Los Angeles, U.S.A.

University of California, Riverside, Riverside, U.S.A.

University of California, San Diego, La Jolla, U.S.A.

University of California, Santa Barbara - Department of Physics, Santa Barbara, U.S.A.

California Institute of Technology, Pasadena, U.S.A.

Carnegie Mellon University, Pittsburgh, U.S.A.

University of Colorado Boulder, Boulder, U.S.A.

– 63 –
Cornell University, Ithaca, U.S.A.

Fermi National Accelerator Laboratory, Batavia, U.S.A.

University of Florida, Gainesville, U.S.A.

Florida International University, Miami, U.S.A.
Y.R. Joshi, S. Linn

Florida State University, Tallahassee, U.S.A.

Florida Institute of Technology, Melbourne, U.S.A.

University of Illinois at Chicago (UIC), Chicago, U.S.A.

The University of Iowa, Iowa City, U.S.A.

Johns Hopkins University, Baltimore, U.S.A.
The University of Kansas, Lawrence, U.S.A.

Kansas State University, Manhattan, U.S.A.

Lawrence Livermore National Laboratory, Livermore, U.S.A.
F. Rebassoo, D. Wright

University of Maryland, College Park, U.S.A.

Massachusetts Institute of Technology, Cambridge, U.S.A.

University of Minnesota, Minneapolis, U.S.A.

University of Mississippi, Oxford, U.S.A.
J.G. Acosta, S. Oliveros

University of Nebraska-Lincoln, Lincoln, U.S.A.

State University of New York at Buffalo, Buffalo, U.S.A.

Northeastern University, Boston, U.S.A.

Northwestern University, Evanston, U.S.A.

University of Notre Dame, Notre Dame, U.S.A.
The Ohio State University, Columbus, U.S.A.

Princeton University, Princeton, U.S.A.

University of Puerto Rico, Mayaguez, U.S.A.
S. Malik, S. Norberg

Purdue University, West Lafayette, U.S.A.

Purdue University Northwest, Hammond, U.S.A.
T. Cheng, J. Dolen, N. Parashar

Rice University, Houston, U.S.A.

University of Rochester, Rochester, U.S.A.
A. Bodek, P. de Barbaro, R. Demina, Y.t. Duh, J.L. Dulemba, C. Fallon, T. Ferbel, M. Galanti, A. Garcia-Bellido, J. Han, O. Hindrichs, A. Khukhunaishvili, P. Tan, R. Taus

Rutgers, The State University of New Jersey, Piscataway, U.S.A.

University of Tennessee, Knoxville, U.S.A.
A.G. Delannoy, J. Heideman, G. Riley, S. Spanier

Texas A&M University, College Station, U.S.A.

Texas Tech University, Lubbock, U.S.A.

Vanderbilt University, Nashville, U.S.A.
University of Virginia, Charlottesville, U.S.A.
M.W. Arenton, P. Barria, B. Cox, R. Hirosky, M. Joyce, A. Ledovskoy, H. Li, C. Neu, T. Sinthuprasith, Y. Wang, E. Wolfe, F. Xia

Wayne State University, Detroit, U.S.A.

University of Wisconsin - Madison, Madison, WI, U.S.A.

†: Deceased
1: Also at Vienna University of Technology, Vienna, Austria
2: Also at IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
3: Also at Universidade Estadual de Campinas, Campinas, Brazil
4: Also at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
5: Also at Université Libre de Bruxelles, Bruxelles, Belgium
6: Also at University of Chinese Academy of Sciences, Beijing, China
7: Also at Institute for Theoretical and Experimental Physics, Moscow, Russia
8: Also at Joint Institute for Nuclear Research, Dubna, Russia
9: Now at Helwan University, Cairo, Egypt
10: Also at Zewail City of Science and Technology, Zewail, Egypt
11: Also at British University in Egypt, Cairo, Egypt
12: Now at Ain Shams University, Cairo, Egypt
13: Also at Department of Physics, King Abdulaziz University, Jeddah, Saudi Arabia
14: Also at Université de Haute Alsace, Mulhouse, France
15: Also at Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
16: Also at Tbilisi State University, Tbilisi, Georgia
17: Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland
18: Also at RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
19: Also at University of Hamburg, Hamburg, Germany
20: Also at Brandenburg University of Technology, Cottbus, Germany
21: Also at MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
22: Also at Institute of Nuclear Research ATOMKI, Debrecen, Hungary
23: Also at Institute of Physics, University of Debrecen, Debrecen, Hungary
24: Also at Indian Institute of Technology Bhubaneswar, Bhubaneswar, India
25: Also at Institute of Physics, Bhubaneswar, India
26: Also at Shoolini University, Solan, India
27: Also at University of Visva-Bharati, Santiniketan, India
28: Also at Isfahan University of Technology, Isfahan, Iran
29: Also at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran
30: Also at Università degli Studi di Siena, Siena, Italy
31: Also at Kyunghee University, Seoul, Korea
32: Also at International Islamic University of Malaysia, Kuala Lumpur, Malaysia
33: Also at Malaysian Nuclear Agency, MOSTI, Kajang, Malaysia
34: Also at Consejo Nacional de Ciencia y Tecnología, Mexico city, Mexico
35: Also at Warsaw University of Technology, Institute of Electronic Systems, Warsaw, Poland
36: Also at Institute for Nuclear Research, Moscow, Russia
37: Now at National Research Nuclear University ‘Moscow Engineering Physics Institute’ (MEPhI), Moscow, Russia
38: Also at St. Petersburg State Polytechnical University, St. Petersburg, Russia
39: Also at University of Florida, Gainesville, U.S.A.
40: Also at P.N. Lebedev Physical Institute, Moscow, Russia
41: Also at California Institute of Technology, Pasadena, U.S.A.
42: Also at INFN Sezione di Padova\textsuperscript{a}, Università di Padova\textsuperscript{b}, Università di Trento (Trento)\textsuperscript{c}, Padova, Italy
43: Also at Budker Institute of Nuclear Physics, Novosibirsk, Russia
44: Also at Faculty of Physics, University of Belgrade, Belgrade, Serbia
45: Also at INFN Sezione di Pavia\textsuperscript{a}, Università di Pavia\textsuperscript{b}, Pavia, Italy
46: Also at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
47: Also at Scuola Normale e Sezione dell’INFN, Pisa, Italy
48: Also at National and Kapodistrian University of Athens, Athens, Greece
49: Also at Riga Technical University, Riga, Latvia
50: Also at Universität Zürich, Zurich, Switzerland
51: Also at Stefan Meyer Institute for Subatomic Physics (SMI), Vienna, Austria
52: Also at Adiyaman University, Adiyaman, Turkey
53: Also at Istanbul Aydin University, Istanbul, Turkey
54: Also at Mersin University, Mersin, Turkey
55: Also at Piri Reis University, Istanbul, Turkey
56: Also at Gaziosmanpasa University, Tokat, Turkey
57: Also at Ozyegin University, Istanbul, Turkey
58: Also at Izmir Institute of Technology, Izmir, Turkey
59: Also at Marmara University, Istanbul, Turkey
60: Also at Kafkas University, Kars, Turkey
61: Also at Istanbul University, Faculty of Science, Istanbul, Turkey
62: Also at Istanbul Bilgi University, Istanbul, Turkey
63: Also at Hacettepe University, Ankara, Turkey
64: Also at Rutherford Appleton Laboratory, Didcot, United Kingdom
65: Also at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom
66: Also at Monash University, Faculty of Science, Clayton, Australia
67: Also at Bethel University, St. Paul, U.S.A.
68: Also at Karamanoğlu Mehmetbey University, Karaman, Turkey
69: Also at Utah Valley University, Orem, U.S.A.
70: Also at Purdue University, West Lafayette, U.S.A.
71: Also at Beykent University, Istanbul, Turkey
72: Also at Bingol University, Bingol, Turkey
73: Also at Sinop University, Sinop, Turkey
74: Also at Mimar Sinan University, Istanbul, Istanbul, Turkey
75: Also at Texas A&M University at Qatar, Doha, Qatar
76: Also at Kyungpook National University, Daegu, Korea