The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

03 November 2014 (v2, 10 November 2014)

Identification of hadronic tau decays in CMS

Christian Veelken for the CMS Collaboration

Abstract

The algorithm used for reconstruction and identification of hadronic tau decays by the CMS experiment at the LHC is presented. The tau reconstruction in CMS takes advantage of the particle-flow algorithm which allows to reconstruct individual hadronic decay modes. The performance of the algorithm in terms of tau identification efficiency and in terms of rates for jets to be misidentified as hadronic tau decays is studied in pp collision data recorded in 2012 at a center-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 19.7 fb^{-1} .

Presented at *Tau14 The 13th International Workshop on Tau Lepton Physics*

Identification of hadronic tau decays in CMS

Christian Veelken^a on behalf of the CMS Collaboration

^a*Laboratoire Leprince-Ringuet, École Polytechnique, 91128 Palaiseau, France*

Abstract

The algorithm used for reconstruction and identification of hadronic tau decays by the CMS experiment at the LHC is presented. The tau reconstruction in CMS takes advantage of the particle-flow algorithm which allows to reconstruct individual hadronic decay modes. The performance of the algorithm in terms of tau identification efficiency and in terms of rates for jets to be misidentified as hadronic tau decays is studied in pp collision data recorded in 2012 at a center-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 19.7 fb^{-1} .

Keywords: CMS, Tau, Reconstruction, Identification, Algorithm

1. Introduction

Tau leptons constitute an important experimental signature for physics analyses at the LHC, such as the analysis of standard model (SM) Higgs $\rightarrow \tau\tau$ production [1] searches for supersymmetry, for leptoquarks, Z' and W' bosons and for non-SM Higgs bosons.

2. CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 Tesla. Within the superconducting solenoid volume are 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. The silicon tracker measures charged particles within the pseudorapidity range $|\eta| < 2.5$. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors.

A particle-flow (PF) algorithm [2, 3] is used to combine information from all CMS subdetectors to identify and reconstruct individual particles in the event, namely

muons, electrons, photons, charged hadrons, and neutral hadrons. The resulting particles are used to reconstruct jets, hadronically decaying tau leptons, and the missing transverse energy vector $\vec{E}_T^{\text{miss}}$, defined as the negative of the vector sum of the transverse momenta of all reconstructed particles, and its magnitude $\cancel{E}_T$.

3. HPS algorithm

Hadronic decays of tau leptons are reconstructed and identified using the “Hadron plus Strips” (HPS) algorithm [4]. The algorithm is designed to reconstruct the individual hadronic tau decays modes $\tau^- \rightarrow \pi^- \nu_\tau$, $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$, $\tau^- \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ and $\tau^- \rightarrow \pi^- \pi^0 \pi^0 \nu_\tau$, using as input the particles reconstructed by the PF algorithm and identified as charged hadrons, photons or electrons. Electrons are considered as input to account for the high probability for photons originating from $\pi^0 \rightarrow \gamma\gamma$ decays to convert to e^+e^- pairs within the volume of the CMS tracking detector. The HPS algorithm is seeded by jets reconstructed using the anti- k_T algorithm [5] with a distance parameter $R = 0.5$.

Neutral pions are reconstructed by clustering the electron/photon (e/γ) constituents of the jet that seeds the tau reconstruction into η - ϕ “strips”. The electrons and photons entering the strips are required to pass

$p_T > 0.5$ GeV. The size of the strips, 0.05×0.20 , is enlarged in ϕ -direction to account for the bending of e^+e^- pairs produced by photon conversions in the 3.8 Tesla magnetic field. Strips containing one or more e/γ constituents and passing a cut of $p_T > 2.5$ GeV on the transverse momentum sum of electrons/photons included in the strip are kept for further processing.

The HPS algorithm builds combinations of one charged hadron plus up to two strips and of three charged hadrons, corresponding to the different hadronic tau decay modes targeted by the algorithm. Charged hadrons are required to satisfy $p_T > 0.5$ GeV. Multiple hadronic tau (τ_h) candidates, given by different combinations of charged hadrons plus strips, are built in parallel. The energy (momentum) of the τ_h candidate is determined by the energy (momentum) sum of charged hadrons plus strips used to build the candidate. Mass window cuts are applied to remove τ_h candidates incompatible with a hadronic tau decay. Hadronic tau candidates are further required to satisfy a collimation criterion: The maximal distance between the τ_h candidate direction and any charged hadron or strip used to build the candidate must not exceed $\Delta R = 3.0/p_T$ [GeV]. The size of the *signal cone* is reduced as function of the candidate p_T , as the collimation of the tau decay products is expected to increase with the Lorentz boost in tau direction. In case $\Delta R = 3.0/p_T$ [GeV] is lower than 0.05 (exceeds 0.10) a signal cone of size $\Delta R = 0.05$ ($\Delta R = 0.10$) is used. The tracks associated to τ_h candidates constructed of three charged hadrons are required to sum to unit charge. In case more than one τ_h candidate passes the mass window, collimation and charge requirements, the τ_h candidate of highest p_T is selected, resulting in a unique hadronic tau candidate to be associated to each jet.

The performance of the tau decay mode reconstruction is illustrated in Fig. 1. The figure shows the mass of τ_h candidates reconstructed in an event sample dominated by $Z/\gamma^* \rightarrow \tau\tau$ production. Hadronic taus reconstructed in different decay modes are represented by different colors. The probability to reconstruct the correct decay mode is higher than 80%. The data is seen to agree well with the expectation for the $Z/\gamma^* \rightarrow \tau\tau$ signal plus backgrounds. The mass of τ_h candidates reconstructed in the decay modes $\tau^- \rightarrow \pi^- \pi^0 \nu_\tau$ and $\tau^- \rightarrow \pi^- \pi^0 \pi^0 \nu_\tau$ (“ $1\pi^\pm + \text{photons}$ ”) is concentrated around the $\rho(770)$ meson resonance, while the mass of τ_h candidates reconstructed in the $\tau^- \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ decay mode (“ $3\pi^\pm$ ”) is aggregated near the $a_1(1260)$ resonance.

Typical acceptance cuts for hadronic taus, applied by CMS physics analyses during LHC run 1, amount to $p_T > 20$ GeV and $|\eta| < 2.3$. Taus in this p_T range pre-

Figure 1: Distribution of τ_h candidate mass in an event sample dominated by $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\tau_h$ events.

dominantly originate from of electroweak bosons, $W \rightarrow \tau\nu$ and $Z/\gamma^* \rightarrow \tau\tau$, for which the cross-section times branching fraction amounts to $O(10 \text{ nb})$ and $O(1 \text{ nb})$ respectively. The taus from W and Z boson decays are used to validate the tau reconstruction primarily. Interesting signal processes, such as SM Higgs $\rightarrow \tau\tau$, are produced with rates of order $O(1 \text{ pb})$. Searches for rare and exotic signals of new physics may be based on cross-sections as low as a few femtobarn. In contrast, quark and gluon jets are produced with cross-sections of $O(100 \mu\text{b})$ at the LHC. The main handle to discriminate hadronic tau decays from the large quark and gluon jet background is the isolation of the tau.

The challenge for using particle isolation at the LHC is the large number of proton-proton interactions occurring in the same bunch crossing, referred to as pileup. On average 9 (21) inelastic pp interactions occurred per LHC bunch crossing in 2011 (2012). Sophisticated techniques have been developed to avoid that particles produced in pileup interactions spoil the isolation of hadronic taus.

The tau isolation is computed by summing the transverse momenta of charged particles of $p_T > 0.5$ GeV plus photons of $E_T > 0.5$ GeV, reconstructed by the PF algorithm, within an *isolation cone* of size $\Delta R = 0.5$, centered on the τ_h direction. The tracks of charged particles considered for the isolation are required to have a distance to the τ_h production vertex, taken as the vertex closest to the highest p_T charged hadron used to build the τ_h candidate, of $\Delta z < 0.2$ cm. Requiring the

charged particles considered for the isolation to be compatible with the τ_h production vertex greatly reduces the effect of pileup. The effect of pileup on the photon isolation is compensated on statistical basis, via the so-called $\Delta\beta$ corrections [6]. The method is based on computing the p_T -sum of charged particles that fail the $\Delta z < 0.2$ cm cut and are within a cone of size $\Delta R = 0.8$ around the tau direction. This p_T -sum is multiplied by a scale factor, which has been determined to make the tau identification efficiency independent of pileup, and subtracted from the photon isolation. Thresholds of 2.0 GeV, 1.0 GeV and 0.8 GeV on the sum of charged isolation plus pileup-corrected neutral isolation define loose, medium and tight working points (WPs) of the tau isolation discriminator. The WPs satisfy different needs of physics analyses in terms of tau identification efficiency versus jet $\rightarrow \tau_h$ misidentification rate.

An MVA-based tau identification discriminator has been trained to improve the separation of hadronic taus from quark and gluon jets. The algorithm uses as input the charged and neutral isolation of the tau plus variables that are sensitive to the lifetime of the tau lepton. The transverse impact parameter of the highest p_T charged hadron used to build the τ_h candidate, defined as the distance of closest approach, in the plane transverse to the beamline, of the track to the τ_h production vertex is used as input for all τ_h candidates. In case of τ_h candidates reconstructed in the $\tau^- \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ decay mode, a fit of the three tracks to a common vertex is attempted. In case the fit succeeds, its distance to the τ_h production vertex is taken as measurement of the tau decay length and used as input to the MVA too. The MVA-based tau identification discriminator reduces the rate of jet $\rightarrow \tau_h$ fakes by 40–50% (for the same tau identification efficiency) compared to the cut based tau isolation discriminator. Very loose, loose, medium and tight WPs of the MVA-based tau identification discriminator are defined by different cuts on the MVA output. The cut thresholds are adjusted as function of τ_h candidate p_T such that the tau identification efficiency for each WP is constant as function of p_T .

Additional discriminators have been developed to separate hadronic tau decays from electrons and muons. An MVA-based algorithm is utilized to discriminate hadronic taus from electrons, using as inputs the compactness and shape of energy deposits in the ECAL, the sum of energy deposits in the HCAL plus the matching between track momenta and the energy deposits in the ECAL. The discriminator against electrons achieves an $e \rightarrow \tau_h$ fake-rate of 0.5% (0.1%) for an efficiency of 90% (70%), in case a loose (tight) cut on the MVA output is applied. The rejection of $\mu \rightarrow \tau_h$ fakes is based on

hits and track segments reconstructed near the direction of the τ_h candidate in the muon system and on the sum of energy deposits in ECAL plus HCAL. The misidentification rate of muons amounts to about 0.1%, for an efficiency of 95%.

4. Validation

The hadronic tau identification efficiency has been measured using $Z/\gamma^* \rightarrow \tau\tau$ events. The measurement is based on selecting a sample of $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\tau_h$ events with no tau identification criteria applied and determining the number of hadronic taus passing (*Pass* region) and failing (*Fail* region) the tau identification discriminator under study. The contribution of backgrounds is sizeable, in particular in the *Fail* region, and is determined by fitting the distributions of the number of tracks, N_{tracks} , that are within a cone of size $\Delta R = 0.5$ around the τ_h direction with shape templates for the $Z/\gamma^* \rightarrow \tau\tau$ signal and for different background processes. The results of the measurement are shown in Fig. 2. The measured tau identification efficiency is in agreement with the Monte Carlo simulation within the combined statistical plus systematic uncertainty of the measurement, typically amounting to 4.5%. The fit has been repeated using the visible mass of muon plus hadronic tau candidate, m_{vis} , as a cross-check. The results obtained using N_{tracks} and m_{vis} agree within uncertainties.

The misidentification rate for jets has been measured in QCD multijet events. The events have been recorded using a single jet trigger with a p_T threshold of 320 GeV. The jet that passes the trigger requirement is excluded from the measurement in order to avoid potential biases of the measured jet $\rightarrow \tau_h$ misidentification rates by trigger requirements. The results of the measurement are shown in Fig. 3. The fake-rates vary by about one order of magnitude as function of jet p_T and range from $O(1\%)$ for low p_T jets and in case loose tau identification criteria are applied to $O(10^{-4})$ for high p_T jets and tight tau identification criteria. The jet $\rightarrow \tau_h$ misidentification rates measured in data agree with the Monte Carlo expectation within 20%. A slight trend is visible for the misidentification rates measured in data to exceed the simulation at low p_T and to fall below the simulation at high p_T .

The discriminators against electrons and muons have been validated using $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ events.

Figure 2: Tau identification efficiency measured as function of τ_h candidate p_T in $Z/\gamma^* \rightarrow \tau\tau$ events, for the cut based tau isolation discriminator (top) and for the MVA-based tau identification discriminator that includes tau lifetime information (bottom), compared to the Monte Carlo simulation.

5. Summary

The HPS algorithm allows to reconstruct individual hadronic decay modes of the tau. Isolation is used as the main handle to suppress the large background of quark and gluon jets. The algorithm has been validated with data and performed well during LHC run 1.

References

- [1] S. Chatrchyan, et al., Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons, JHEP 1405 (2014) 104, arXiv:1401.5041, doi:10.1007/JHEP05(2014)104.
- [2] CMS Collaboration, Particle-flow event reconstruction in CMS and performance for jets, taus, and E_T , CMS Physics Analysis

Figure 3: Jet $\rightarrow \tau_h$ misidentification rates measured in QCD multijet events, for the cut based tau isolation discriminator (top) and for the MVA-based tau identification discriminator (bottom).

Summary CMS-PAS-PFT-09-001 (2009).

URL <http://cdsweb.cern.ch/record/1194487>

- [3] CMS Collaboration, Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector, CMS Physics Analysis Summary CMS-PAS-PFT-10-001 (2010). URL <http://cdsweb.cern.ch/record/1247373>
- [4] Performance of tau-lepton reconstruction and identification in CMS, JINST 7 (2012) P01001. arXiv:1109.6034, doi:10.1088/1748-0221/7/01/P01001.
- [5] M. Cacciari, G. P. Salam, G. Soyez, The anti-kt jet clustering algorithm, JHEP 04 (2008) 063. arXiv:0802.1189.
- [6] CMS Collaboration, Tau id performance plots, CMS Detector Performance Note CMS-DP-2014-015 (2014). URL <https://cds.cern.ch/record/1704439>
- [7] S. Chatrchyan, et al., The CMS experiment at the CERN LHC, JINST 3 (2008) S08004. doi:10.1088/1748-0221/3/08/S08004.