TAU IDENTIFICATION AND RECONSTRUCTION AT THE CMS*

Robert Boniecki

University of Warsaw, Faculty of Physics
Hoża 69, 00-681 Warszawa, Poland

(Received April 16, 2013)

A dedicated tau reconstruction algorithm developed by the CMS Collaboration is presented in detail with results of the reconstruction performance analysis. Two different tau isolation methods used by the algorithm are discussed and compared.

DOI:10.5506/APhysPolB.44.1379
PACS numbers: 14.60.–z, 13.35.–r

1. Introduction

Tau leptons play a significant role in a wide range of physics analyses. Efficient and robust tau reconstruction is essential for searches for the Higgs boson, searches for the supersymmetric particles or precise electroweak measurements. The tau is the heaviest lepton and its lifetime is too short to reach any subdetector, therefore, it has to be reconstructed indirectly from its decay products. The reconstruction process requires sophisticated algorithms that take advantage of distinct decay modes and intermediate state constraints.

1.1. Tau lepton decay modes

Approximately two out of three tau leptons decay hadronically. Most of these hadronic decays are mediated by a resonant state, which can be used to implement additional constraints to the decay mode reconstruction algorithm. The main tau lepton decay modes with corresponding branching ratios [1] and possible intermediate state information are shown in Table I.

* Presented at the Cracow Epiphany Conference on the Physics After the First Phase of the LHC, Kraków, Poland, January 7–9, 2013.

TABLE I

Main tau lepton decay modes, $h^{\pm}$ stands for charged hadron, namely pion or kaon.

Decay mode	Resonance	Mass [MeV/c^2]	Branching ratio [%]
$\tau^- \to h^- \nu_\tau$			11.6
$\tau^- \to h^- \pi^0 \nu_\tau$	ρ^-	770	26.0
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	a_1^-	1200	10.8
$\tau^- \to h^- h^+ h^- \nu_\tau$	a_1^-	1200	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$			4.8
Other hadronic			1.7
Total hadronic			64.8
$\tau^- \to e^- \overline{\nu}_e \nu_\tau$			17.8
$\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$			17.4
Total leptonic			35.2

1.2. Experimental signature

The experimental signature for hadronic tau decays is an isolated and collimated jet with low charged track multiplicity. The number of charged particles forming tau candidate is usually restricted only to one or three, which makes a powerful discriminant against quark or gluon jets which tend to have more charged constituents. In the case of leptonic tau decays, the signature is a single, isolated lepton. It is important to note that at least one neutrino is present in all of the decays, which gives rise to missing transverse energy and does not allow to reconstruct full tau mass.

2. CMS tau reconstruction algorithm

The tau reconstruction at the CMS [2] experiment is seeded by jets, in search for tau hadronic decay modes signatures. To achieve the best possible performance, all available information is used. Consequently, the CMS tau reconstruction algorithm takes advantage of the decay mode reconstruction with intermediate resonant state constraints, collimation of the jet, tau mass constraint and appropriate isolation criteria. The algorithm used for hadronic tau reconstruction at CMS is called Hadron Plus Strips (HPS).

2.1. Particle Flow

The HPS algorithms uses Particle Flow (PF) technology — a PF jet object is used as a starting point for HPS. Particle Flow is an sophisticated algorithm used for combination of the information provided by every sub-detector of the CMS detector, in order to reconstruct physics objects (particle candidates). Particle Flow provides particles of following types: photon, charged hadron, neutral hadron, muon, electron. The higher level objects, like jets or missing transverse energy, are reconstructed using the particle candidates.

2.2. Hadron Plus Strips algorithm

Starting from the Particle Flow jet as an input, the HPS algorithm uses PF charged hadron objects and independently reconstructs π^0s from PF electrons and photons, as objects called strips. Strips are defined as regions of space in ECAL of size $\Delta\eta = 0.05$, $\Delta\phi = 0.2$ corresponding to the π^0 signal. Neutral pions almost immediately decay into two photons and those can convert to electron–positron pairs, which are bent in the magnetic field, thus causing broadening of the signal in azimuthal direction. This effect is taken into account by defining the strip to be wider in ϕ than in η direction. For accurate reconstruction, the mass of the strip object has to be consistent with π^0 mass. After the decay mode identification, the charged tau decay products have to fall inside the so-called shrinking cone, defined as a cone of size that gets lower for higher p_{T}: $\Delta R = (2.8\ \mathrm{GeV}/c)/p_{\mathrm{T}}^{\tau}$. Additionally, the direction of the reconstructed tau has to match the direction of the original PF jet, with deviation lower than $\Delta R = 0.1$.

2.3. Decay mode reconstruction

HPS algorithm uses four different topologies to reconstruct individual tau hadronic decay modes:

1. **One charged hadron** — reconstructs $\tau^- \to h^-\nu_\tau$ decays and $\tau^- \to h^-\pi^0\nu_\tau$ decays when the neutral pion energy is too low for strip reconstruction.

2. **One charged hadron + one strip** — reconstruction of $\tau^- \to h^-\pi^0\nu_\tau$ when the photons from π^0 decay are collimated in one strip.

3. **One charged hadron + two strips** — aimed at the reconstruction of $\tau^- \to h^-\pi^0\nu_\tau$ in events in which the photons from π^0 decay are separated in two distinct strips.

4. **Three charged hadrons** — corresponding the $\tau^- \to h^-h^+h^-\nu_\tau$ decay. Presence of three tracks allows to put additional constraints for a common secondary vertex of those tracks.

The decay modes are required to be consistent with corresponding intermediate state hypothesis (ρ, a_1). In the case more than one hypothesis is compatible with the decay, the one with the highest p_{T} of the reconstructed tau is given preference.

3. Isolation

To reduce the number of quark or gluon jets misidentified as a tau, the sum of energy deposits in the jet not assigned to the tau has to be appropriately small. The CMS Collaboration uses two different methods — both well established and thoroughly validated — simpler, cut-based approach and multivariate method, which shows superior efficiency and fake-rate.

3.1. Cut-based isolation

The cut-based isolation defines an isolation cone and sums p_{T} of all charged hadrons and photons inside isolation cone that do not contribute to the tau signal. Three different working points for this isolation are defined: Loose, Medium and Tight. An important aspect of this method is to provide a solution to the pile-up, which greatly affects the sum of p_{T}. Charged particles coming from pile-up can be distinguished by a vertex constraint, however, neutral pile-up needs to be estimated. The estimation of neutral pile-up component comes from charged pile-up tracks inside the isolation cone and takes advantage of an experimental fact that neutral pile-up is proportional to charged pile-up.

3.2. Multivariate isolation

The multivariate (MVA) isolation is based on energy deposits in five rings around the tau. The improvement upon the cut-based method comes from the information of radial distribution of p_{T} inside the isolation cone. MVA method uses a Boosted Decision Tree trained against misidentification of non-tau jets as hadronic tau decays.

4. Results

4.1. Reconstruction efficiency

The reconstruction efficiency is calculated by processing a $Z^0 \to \tau^-\tau^+$ MC sample with the HPS algorithm. The efficiency is defined as a ratio of the number of reconstructed taus to the number of generated taus. Figure 1 shows the dependence of reconstruction efficiency on the generated visible tau p_{T} for both cut-based and MVA isolation methods. The efficiency is essentially flat for p_{T} higher than 25 GeV for both isolation methods. The MVA shows superior efficiency, reaching 70% for the Loose isolation working point.

The efficiency as a function of number of reconstructed vertices is presented in Fig. 2. The dependence of the efficiency on pile-up is weak.

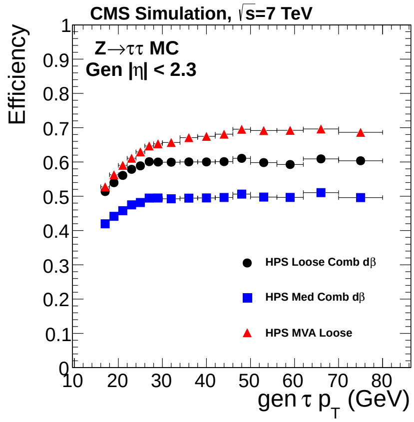

Fig. 1. Reconstruction efficiency as a function of generated visible tau p_{T} for $Z \to \tau^-\tau^+$ MC sample. Two isolation working points for cut-based method: Loose (black circles) and Medium (blue squares), as well as MVA Loose working point (red triangles) are shown.

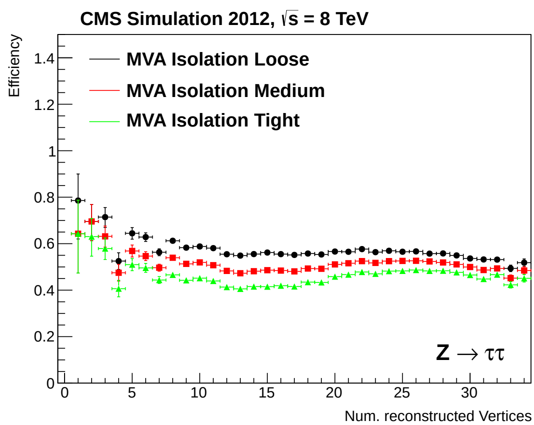

Fig. 2. Reconstruction efficiency as a function of number of reconstructed vertices for MVA isolation for Loose, Medium and Tight working points.

4.2. Fake-rate

The purity of the HPS reconstruction is studied in terms of fake-rate, defined as probability of a quark or gluon jet being misidentified as tau. Fake-rate is estimated from data, using events enriched in quark or gluon jets, namely $W+$jets events. Results presented in this contributions were obtained with 3 fb^{-1} of data collected by the CMS detector in early 2012. The fake-rate as a function of jet p_{T} are shown in Fig. 3. The MVA method provides 10–20% better fake-rate than the cut-based isolation for low p_{T} region.

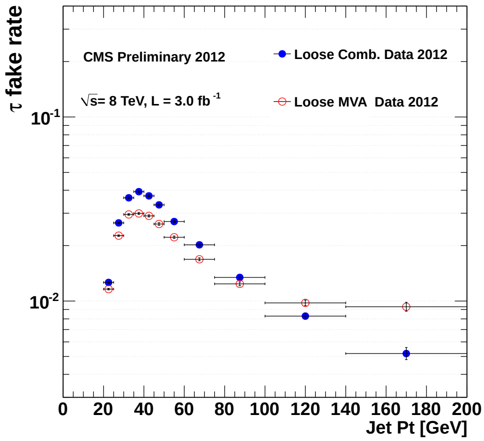

Fig. 3. Fake-rate as a function of jet p_{T} for both cut-based and MVA isolations. The fake-rate is estimated from $W+$jets events selected from 2012 data.

5. Conclusions

Tau reconstruction and identification at the CMS is made using a dedicated algorithm called Hadron Plus Strips, which takes advantage of different tau decay modes. Two different methods of isolating the tau are used, cut-based approach and multivariate analysis. Both provide very good efficiency and purity of tau reconstruction at the CMS.

The author was supported by the HOMING PLUS programme of the Foundation for Polish Science, cofinanced from the European Union, Regional Development Fund.

REFERENCES

[1] J. Beringer *et al.* [Particle Data Group], *Phys. Rev.* **D86**, 010001 (2012).
[2] R. Adolphi *et al.*, *JINST* **0803**, S08004 (2008).