Tau Identification Performance at CMS experiment

Presented by,

VINAYA KRISHNAN MB
Institute of Physics, Bhubaneswar

CMS collaboration
INTRODUCTION

- The mass of Tau lepton is $m_\tau = 1776.86 \pm 0.12$ MeV and it is the only lepton that can decay weakly to hadrons (a.k.a Hadronic tau).
- Hadronic taus ($\tau_h$) are reconstructed and identified using hadron plus strip (HPS) algorithm.
- Decay products of $\tau_h$ are mainly combination of charged and neutral pi-mesons and tau nutrino, this makes it more difficult to distinguish $\tau_h$ decay products from quark and gluon jets.
- Two discriminants are used to separate $\tau_h$ from jets. Such as,
  - Isolation sum discriminant.
  - Multi variant analysis (MVA) based discriminant.
- To measure $\tau_h$ isolation and identification efficiencies we used tag and probe method using $Z/\gamma^* \rightarrow \tau\tau \rightarrow \tau_\mu\tau_h$ events.
Start from Anti$K_T$ ($R=0.4$) PF jet
Reconstruct decay modes with one or three charged hadrons, and one or two neutral pions
Neutral Pions reconstructed using an elongated $\eta \times \phi$ strips collecting energy spread from photon conversions due to magnetic field.
Charged hadrons and photons reconstructed from tracks and calorimeter energy using a particle-flow technique
Signal constituents are required to be in a smaller cone ($p_T$ dependent cone, $R_{\text{sig}} = 3\text{GeV}/p_T$)
Mass constraints compatible with $\rho$ and $a_1$ meson mass
The charged pion nuclear interaction with tracker material causes cascades of low $p_T$ electrons and photons, that can appear outside of strip. And photons from neutral pions also have strong probability to convert electron-positron pair, after multiple scattering and bremsstrahlung they may appear in the outside the strip.

Naively, these decay products can be integrated into the strip by suitably increasing its size, called Dynamic strip reconstruction.

The strip size is the function of $e/g p_T$ which can be written as,

$$
\Delta \eta = f(p_T^\gamma) + f(p_T^{\text{strip}})
$$

$$
\Delta \phi = g(p_T^\gamma) + g(p_T^{\text{strip}}).
$$
Isolation Sum Discriminants
(Cut based Isolation)

The isolation sum defined as,

\[ I_\tau = \sum P_T^{h\pm} (d_Z < 2\text{mm}) + \max(0, \sum P_T^{\gamma} - \Delta\beta \sum P_T^{h\pm} (d_Z > 2\text{mm})) \]

Charged isolation  Photon isolation  Pileup correction

**Charged Isolation**: Summing the \( p_T \) of charged particle within the isolation cone (radius = 0.5(0.3)) of centered on the direction of the \( \tau_h \) candidate. The pileup (PU) contribution is suppressed by taking account of charged particle originated within the < 0.2 cm distance from the \( \tau_h \) production vertices.

**Photon Isolation**: Summing the \( p_T \) of photons within the isolation cone (radius = 0.5(0.3)) of centered on the direction of the \( \tau_h \) candidate.

**Pileup correction**: PU corrections for photon isolation estimated from \( p_T \) sum of charged particles of > 0.2 cm distance from the \( \tau_h \) production vertices, but cone of Radius 0.80. The imperical constant \( \Delta\beta \) value is 0.2 in 2016 data.
Cut based Isolation

Dynamic strip reconstruction allows photon candidate outside the signal cone can contribute to the signal, which also effectively increase the jets-$\tau_h$ mis-identification probability. Which introduce an additional quantity called $p_T$ strip outer, it defined as:

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

It is the $p_T$ sum of $e/\gamma$ candidates included in the strip, but located outside of the signal cone. (see the red $\gamma$). Applying a cut on $p_T$ outer less than 10% of $p_T$ tau reduce around 20% of jets-$\tau_h$ mis-identification probability.
MVA based Discriminant

A classifier based on decision boosted trees (BDT) is used to discriminate $\tau_h$ decay product from quark and gluon jets. The variables used as inputs of to the BDT are,

➢ Isolation sum
➢ $\tau_h$ decay modes
➢ The impact parameter $d_0$ of the highst $p_T$ track of $\tau_h$ candidate.
➢ The distance between the $\tau$ production and decay vertices.
➢ $p_T$ strip outer
➢ Multiplicity of photon and electoron, candidates.
Tau Isolation Working points (WP)

- VVLoose, VLoose, Loose, Medium and Tight WP are defined for the cut-off isolation discriminant by isolation sum $I_\tau < 4.5, 3.5, 2.0, 1.0$ and 0.8 respectively.

- The MVA working points are defined on the output of the discriminator for different identification efficiencies: very loose (90%), loose(80%), medium(70%), tight (60%), very tight (50%), very very tight (40%).

\[
\frac{\tau_h \text{ reco+ID efficiency}}{\text{Misidentification probability}} = \frac{\text{denominator} \&\& \tau_h p_T > 20 \text{ GeV} \&\& |\eta| < 2.3 \&\& \text{decay modes} \&\& \text{isolation}}{\text{gen vis } \tau p_T > 20 \text{ GeV} \&\& |\eta| < 2.3 \&\& \text{gen. had. decay modes}}
\]

\[
= \frac{\text{denominator} \&\& \tau_h p_T > 20 \text{ GeV} \&\& |\eta| < 2.3 \&\& \text{decay modes} \&\& \text{isolation}}{\text{gen q/g jet } p_T > 20 \text{ GeV} \&\& |\eta| < 2.3}
\]
• Efficiencies are measured using Tag and Probe technique.
• $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\tau_h$ events are used, where muon ($p_T > 29\text{GeV}$) as a Tag and $\tau_h$ ($p_T > 20\text{GeV}$) is used as probe.
• Invariant mass ($m_{\text{vis}}$) distribution of the muon and $\tau_h$ is fitted to extract the efficiency.
μτₜₜ visible mass for passing probe

Comparision distribution of identified τₜₜ from its decay modes on 2016 and 2017 data.
Measurement of $e/\mu \rightarrow \tau_h$ misID probability

The mis-identification probability is measured similarly by tag and probe method using $Z/\gamma^* \rightarrow ee$ (or $\mu\mu$) events.

![Graph showing the measurement of $e\rightarrow \tau_h$ misID and $\mu\rightarrow \tau_h$ misID]
Distribution of $\tau_h$ mass

Comparision distribution of identified $\tau_h$ from its decay modes on 2016 and 2017 data.
jet→τ_h mis-id probabilities is measured using W(µν) +jets events and tt→eµ + jets events.

**Jets→τ_h Fakerate measurement**

**Tau-16-003**

![Graphs showing jet→τ_h mis-id probabilities as a function of p_T^{jet} for W+jets and e+jets events.](image)
SUMMARY

• Discussed $\tau_h$ identification methods employed at CMS.
• Presented its performance.
• Performance are measured from data in terms of efficiency and mis-id probability.
• The measured performance in data is consistent with that of the simulation.
The unknown functions are determined from a single τ lepton MC event sample, generated with uniform $p_T$ in the range of 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.

$$f(p_T) = 0.20 \cdot p_T^{-0.66}$$

$$g(p_T) = 0.35 \cdot p_T^{-0.71}.$$