Identification of hadronic tau decays in CMS

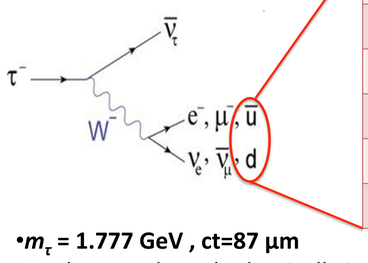
R. Venditti (University and INFN Bari)
on behalf of the CMS Collaboration

ICHEP 2014: 37th International Conference on High Energy Physics, 2-9 July 2014, Valencia

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 ID efficiency, rate of jet→tau, e→tau and mu→tau fakes and in terms of tau energy reconstruction is measured 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.

Importance of tau leptons

Tau leptons constitute an important experimental signature for physics analyses at the LHC:
- Evidence of a SM Higgs Boson decaying into tau leptons[1].
- Searches for MSSM neutral and charged Higgs bosons.
- Drell–Yan, W boson and top quark pair production in final states with τ.
- Searches for supersymmetry, leptoquarks , Z' and W' bosons as well as for doubly charged Higgs bosons.

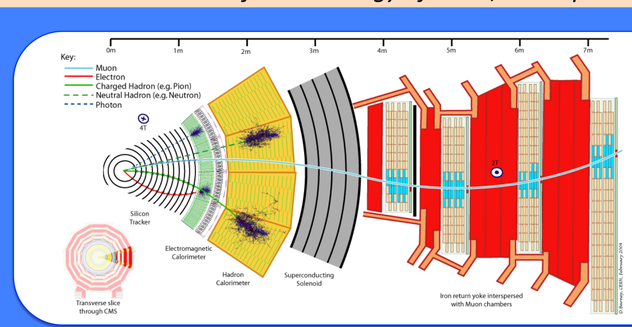


Decay Mode	Resonance	BR
τ→eνν		18%
τ→μνν		17%
τ→hν		12%
τ→hν	ρ(770)	37%
τ→hhhν	a1 (1200)	15%

- m_τ = 1.777 GeV , ct=87 μm
- Tau leptons decay hadronically into either one or three charged mesons (predominantly π^+, π^-) in presence of up to two neutral pions. The latter decay instantaneously via $\pi^0 \rightarrow \gamma\gamma$.
- Muons and electrons from tau decay cannot be distinguished from ones produced in the primary interaction.

The algorithms for tau reconstruction and identification in CMS aim at hadronic tau decays (τ_h).
It benefits from tau decay features, i.e. low multiplicity and high collimation if compared to jet produced in QCD events.

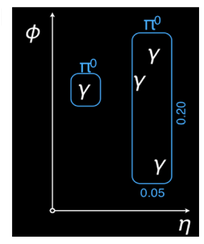
The CMS Experiment



- Superconducting solenoid providing 3.8-Tesla B-field[2].
- Combining information from all the subdetectors→Global event description thanks to the Particle Flow algorithm[3].
- Charged and neutral hadrons, photons, muons, electrons are first identified and then used to reconstruct the missing energy (E^{miss}), jets, τ and lepton isolation.
- Whole data collected in 2012 corresponding to an integrated luminosity of **19.7/fb at 8 TeV** used.

Tau decay mode reconstruction

Hadronic decays of tau leptons are reconstructed and identified by the **Hadrons plus Strips**[5] (HPS) algorithm seeded by jets reconstructed with Particle Flow.

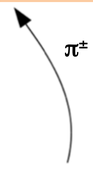


Strip
- Photons from π^0 can be converted into the tracker.
- Electrons from photons conversion are bended from the B-field ,affecting the typical π^0 signature in the η direction.
- Clustering the photons or electrons into η−ϕ strips 0.2x0.05 enlarged in η direction to account for possible broadening of the calorimeter signature by photon conversions

Combining hadrons plus strips

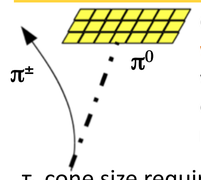
Combinatorial approach is taken for building the combinations of charged particles plus strips that correspond to different hadronic tau decay modes .

Single Hadron: τ→π±



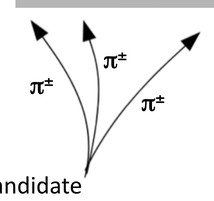
A single charged hadron without any strips.

Hadron + Strip: τ→ ρ→π±π0



Check compatibility with a **ρ mass window**. The size of the mass window is enlarged for high P_T τh candidates to account for resolution effects.

Three Hadrons: τ→ a1→π±π±π±



Combination of three charged hadrons that fits within a **a1 mass hypotesis**. The tracks are required to be compatible with originating from the same event vertex.

τ_h cone size requirement according to the P_T of the candidate

Tau Isolation

Isolation is a key handle in reducing background from quark or gluon jets that can easily fake a τ_h.
Two approaches have been devoleped: cut-based and MVA discriminator.

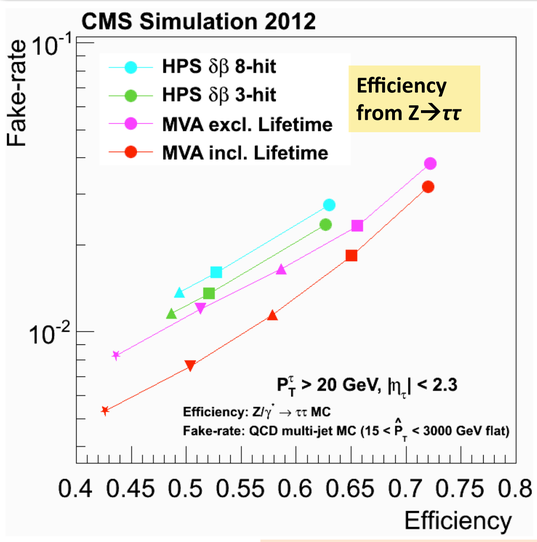
Cut Based Isolation

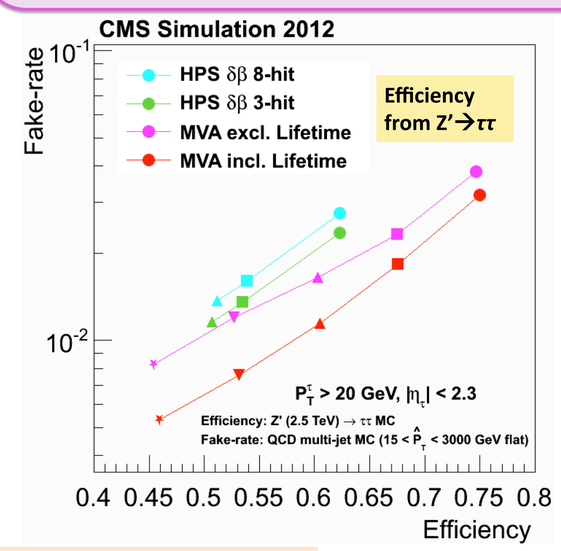
$$I_\tau = \sum P_T^{ch}(dz < 0.2cm) + \max(P_T^\gamma - \Delta\beta, 0)$$

The Δβ correction is computed by summing the transverse momenta of charged particles with $d_z > 0.2$ cm wrt τ_h production vertex and are within a cone of size ΔR = 0.8 around the τ_h direction. The sum is scaled by a factor 0.4576, chosen to make the τ_h identification efficiency insensitive to pile–up.

MVA Isolation

Combine kinematic and τ_h lifetime information. Variables:
- P_T, η, decay mode of the τ_h candidate.
- Isolation variables (ΣP_T^{ch}, P_T^γ,pile-up contribution).
- Transverse impact parameter of the leading track and its significance (1prong).
- The distance between tau production and decay vertex and its significance (3prong).

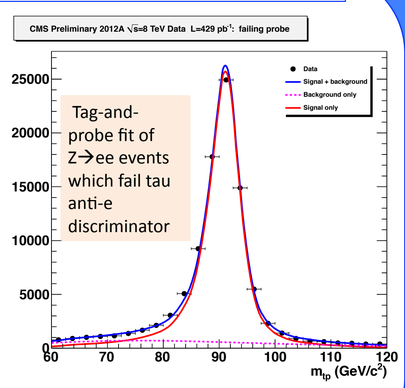




MVA isolation reduces the jet→τ fake rate by a factor 2 compared to the cut-based isolation. Significant improvement for high-pt taus.

Discrimination against electrons

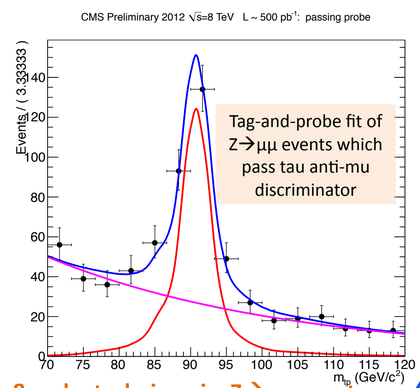
High probability for an electron to be recostructed as 1prong τ_h
Cut-based discriminator: cut is applied on the output of an MVA that has been trained to separate electrons from charged pions.
MVA-based dicriminator: 16 categories based on 4 different reason for an electron to fake a τ_h. 16 BDTs are built to obtain the lowest possible e → τ_h fake rate. Input variables:
- τ variables (P_T,η,electromagnetic energy fraction, E/P, H/P, m_τ)
- Strip variables associated to the photon reconstructed in the strip.
- Track variables: PF electron MVA output for the leading particle, χ2,numbers of hits in silicon pixel plus strip tracking detector, ln(P_T), η.
- Electron variables related to the electron matched to the τ_h.



The fake rate from electrons is measured from data with a tag&probe technique using electrons from Z→ee events.

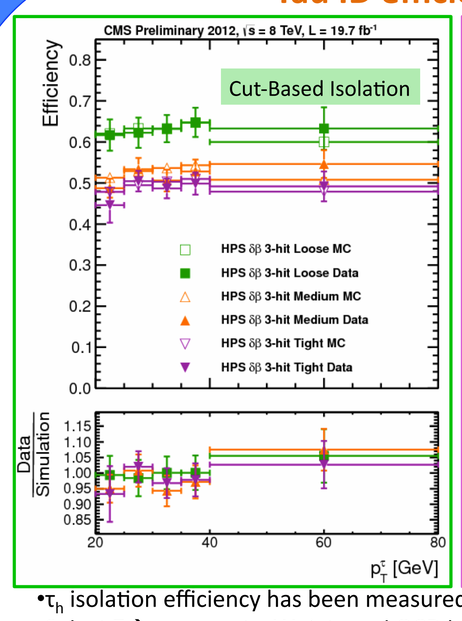
Discrimination against muons

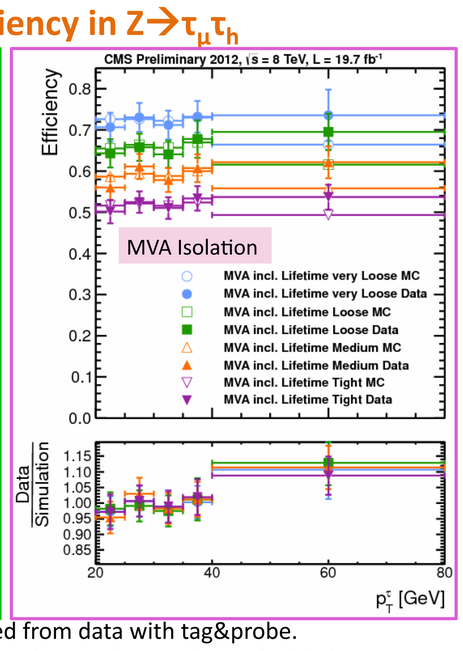
High probability for a muon to be reconstructed as 1prong τ_h.
Cut based anti-μ discriminator: vetoing taus in case signals in the muon system are found near the τ_h direction.
MVA-based anti-μ dicriminator: input variables are:
- ECAL and HCAL energies associated to the charged leading track as well as the ones associated to photons and other τ constituents.
- Fraction of τ_h transverse momentum carried by the leading particle.
- Number of track segments in the muon system reconstructed within a cone of size ΔR = 0.5 around the τ_h direction.
- Number of hits in the muon system detected within a cone of size ΔR = 0.5 centered on the τ_h direction.



The fake rate from muons is measured from data with a tag&probe technique in Z→μμ events.

Tau ID efficiency in $Z \rightarrow \tau_\mu \tau_h$



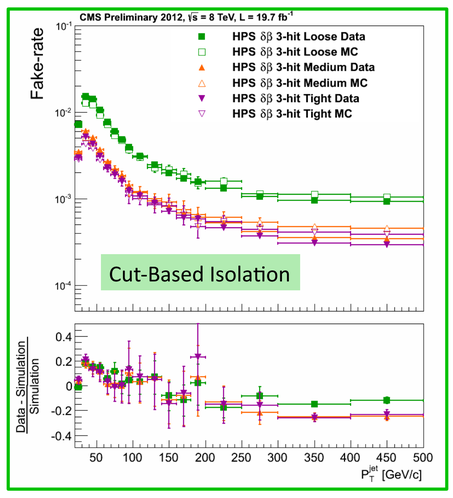


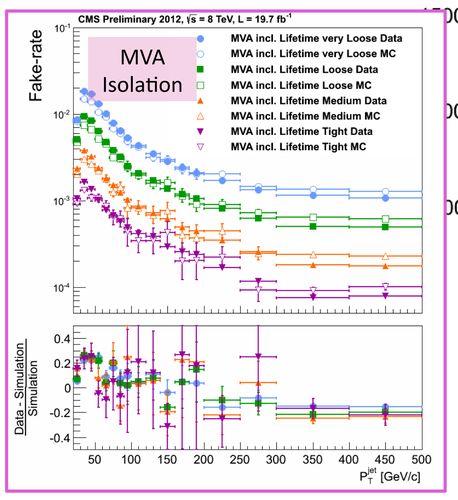
- τ_h isolation efficiency has been measured from data with tag&probe.
- Select $Z \rightarrow \tau_\mu \tau_h$ events. W+jets and QCD background are estimated with data driven technique.
- Simultaneous fit of the Pass and Fail regions using track multiplicity as an observable to perform the fit ($m^{vis}_{\mu\tau}$ also used as a cross-check).
- The fit is performed in τ_h candidate P_T bins using a tool based on Roofit[6] tool, with τ_h identification efficiency being the parameter of interest.
- Systematic uncertainties are represented by nuissance parameters in the fit model.

Data-to-MC scale factors compatible with 1 in each pt bin with a 5% uncertainty.

Jet→τ Fake Rate in QCD events

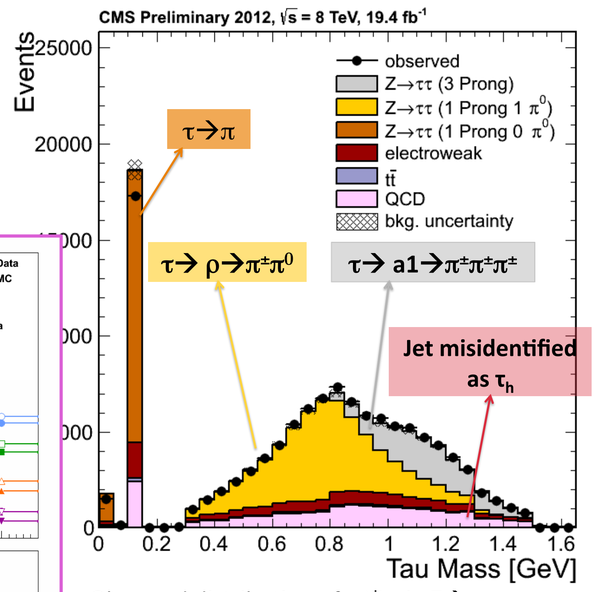
- Rate with which quark or gluon jet pass τ_h isolation discriminators is measured in QCD events for both cut-based and MVA discriminators.
- QCD events selected by requiring at least two jets of P_T > 20 GeV, |η| < 2.5.
- Measurement performed in both MC and data→observed difference varies with jet P_T due to the uncertainties with which MC simulations models the hadronization and fragmentation for high P_T jets.





Fake rate from few per-cent for 50 GeV P_T jet to few per-mille for high P_T jet has been observed.

Tau ID Performance



- Observed distribution of $m^{vis}_{\mu\tau}$ in $Z \rightarrow \tau_\mu \tau_h$ events.
- Full selection described above are applied to the τ_h candidate.
- The Z→ττ contribution is split according to the tau decay mode reconstructed by HPS.
- After a best fit procedure, $m^{vis}_{\mu\tau}$ is used for energy scale calibration.
- Scale factor and energy scale corrections are extracted for each decay mode .

[1]CMS Collaboration, "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons" , CMS-PAS-HIG-13-004, http://arxiv.org/abs/1401.5041
[2]CMS Collaboration, "The CMS experiment at the CERN LHC", JINST03,(2008) S08004
[3]CMS collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus and E^t_{miss}", CMS-PAS-PFT-09-001, 2009.
[4]CMS Collaboration, "Tau-ID Performance Plots", CMS DP-2014/015
[5]CMS Collaboration, "Performance of tau-lepton reconstruction and identification in CMS", *JINST* **7** (2012) P01001, arXiv:1109.6034 .
[6]Higgs PAG, https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit.