



中國科學院高能物理研究所  
*Institute of High Energy Physics*  
*Chinese Academy of Sciences*  
Beijing

ICNFP 2015 - 23 to 30 August, Kolymbari

# Reconstruction and identification of tau leptons in CMS

Francesco Romeo

On behalf of the CMS collaboration

# Motivations for the tau reconstruction and identification

**Tau leptons** constitute an **important experimental signature** for analyses at the **CERN LHC**

- **Standard Model:**

- Drell-Yan, W boson, top-quark pair production in final states with tau leptons

- **Higgs boson:**

- Tau is the heaviest lepton  
→ important to measure the Higgs-Tau coupling
- Searches for MSSM neutral and charged Higgs

- **Beyond Standard Model:**

- Supersymmetry, 3rd-generation leptoquarks,  $Z'$  and  $W'$  bosons
- Some models favour the production of the 3rd-generation fermions  
→ important to characterise the properties of potential new particles

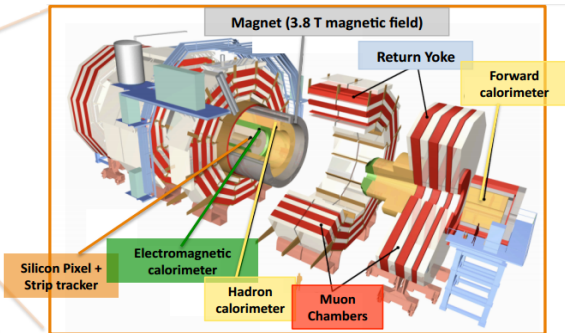
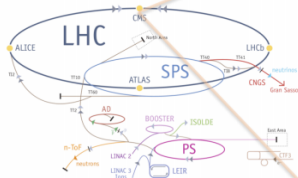
# Tau properties

- **Mass:** 1.78 GeV
- **Lifetime:**  $c\tau = 87 \mu m$
- Decay **branching** ratio (BR)

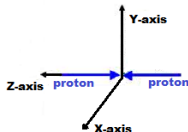
Decay Mode	Resonance	$\mathcal{B}$ [%]
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
$\tau^- \rightarrow h^- \nu_\tau$		11.5
$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770)$	26.0
$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	10.8
$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other hadronic modes		1.8
All hadronic modes		64.8

- The CMS tau algorithm aims to reconstruct **tau leptons decaying hadronically** ( $\tau_h$ )
- Semileptonic tau decays occur about **2/3 of the time** and may end up into **one or three charged** mesons ( $\pi$  or less frequently  $K$ ), called **prongs** ( $h^\pm$ ), in presence of **up to two**  $\pi^0$
- **Topology features:** low multiplicity and high collimation of  $\tau_h$  constituents, non-negligible lifetime compared to jet produced in QCD events
- Presence of neutrino(s) in the final state  $\rightarrow$  **presence of missing transverse energy**

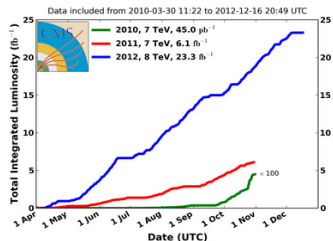
## The CMS detector at LHC



### Cartesian coordinates



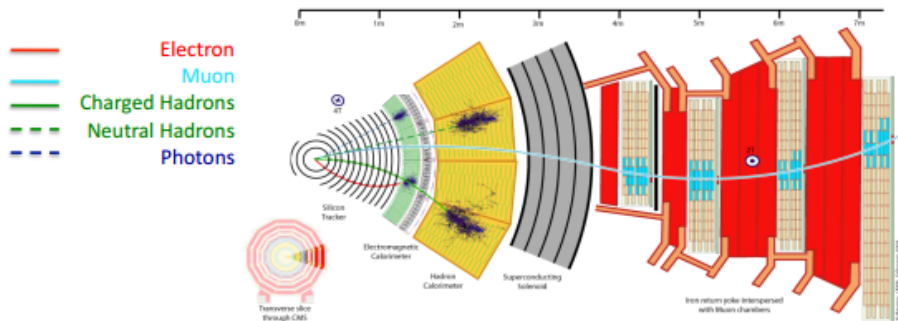
### CMS Integrated Luminosity, pp



# The particle flow algorithm

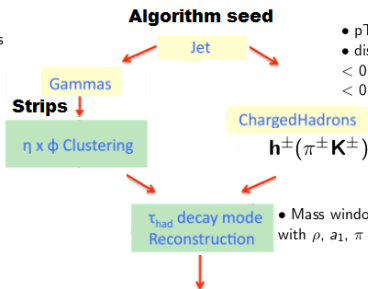
CMS uses a **particle flow** algorithm to provide a global event description

- It combines **information from all subdetectors**
- It **reconstructs muons, electrons, photons, charged and neutral hadrons**
- These objects are **used as input to reconstruct higher-level objects** like: missing energy, lepton isolation, jets, and **hadronic taus**



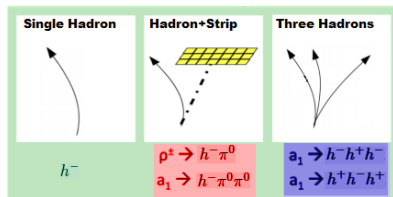
# $\tau_h$ reconstruction in CMS with the Hadron plus Strips (HPS) algorithm

- Photons from  $\pi^0 \rightarrow \gamma\gamma$  decays may convert to  $e^+e^-$  pairs
- Define 0.05 - 0.20 ( $\eta - \phi$  plane) strips to look for  $e, \gamma$  of  $p_T > 0.5$  GeV
- $\pi^0$ : strips with  $\geq 1$   $e, \gamma$  candidates with overall  $p_T > 2.5$  GeV

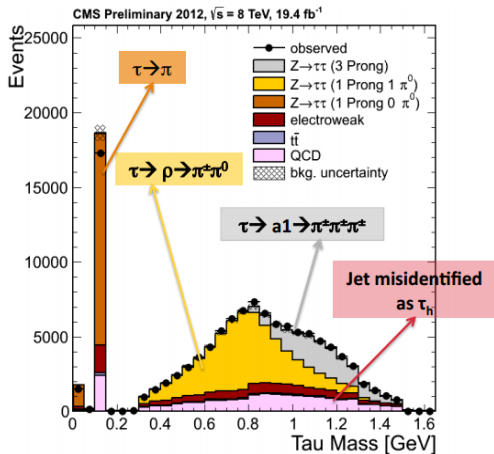


- $p_T > 0.5$  GeV
- distance between  $trk$  and  $\tau_h$  production vertex  $< 0.4$  cm in the direction of the beamline
- $< 0.03$  cm in the transverse plane

Decay channel	BR (%)
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$	17.36
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$	17.85
$\tau^- \rightarrow h^- \nu_\tau$	11.6
$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	26.0
$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	9.5
$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	9.8
$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$	4.8
others	3.1



## Tau decay mode reconstruction

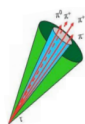


- $\tau_h$  mass reconstructed in  $Z \rightarrow \mu\tau_h$  events
- The distribution is split according to different hadronic tau decay modes
- Good performance observed
- This observable is used to calibrate the energy scale

# Tau identification against quark and gluon jets

- Quark and gluon jets may pass the tau reconstruction criteria of the HPS algorithm
- In order to reduce the rate of jets reconstructed as  $\tau_h$ , the HPS algorithm relies on tau isolation

## Cut-based isolation



Isolation cone  
size:  $\Delta R = 0.5$

$$I = \sum P_T^{ch} (|dz| < 2mm) + \max\left(0, \sum P_T^{neutral} - k \times \sum P_T^{ch} (|dz| > 2mm)\right)$$

Charged Particle inside  
isolation cone from hard  
interaction

Gammas inside  
isolation cone

Factor chosen to make  
the isolation flat wrt # of  
vertices in the event

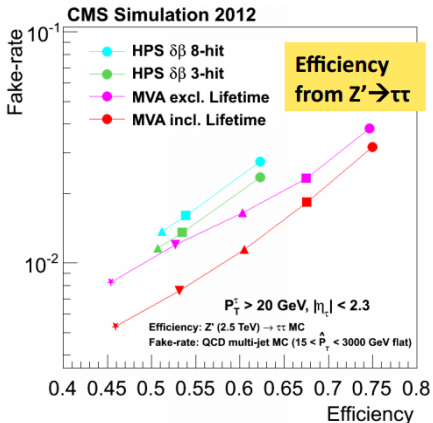
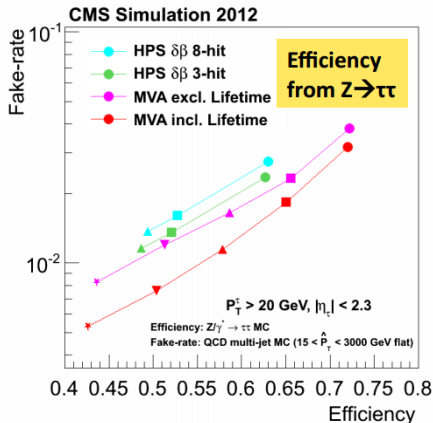
Energy from  
pile-up  
interactions

- $I < 2.0, 1.0, 0.8$  GeV defines 3 working points (Loose, Medium, Tight)

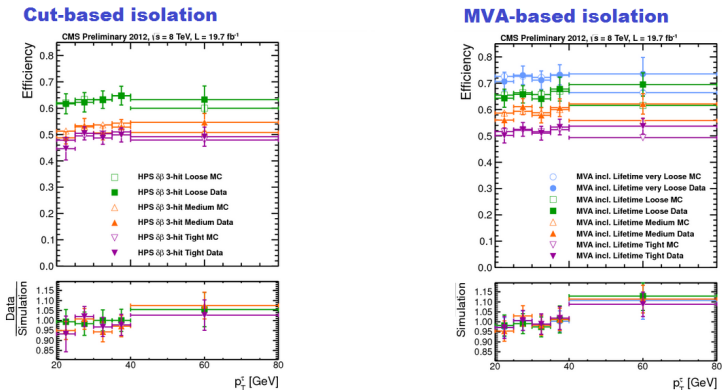
## MVA-based isolation

It combines tau isolation and tau lifetime information

- $p_T, \eta$ , decay mode of the  $\tau_h$  candidate
- All cut-based isolation variables separately
- Transverse impact parameter of the leading track and its significance
- Tau flight distance and its significance (3  $h^\pm$  only)

Expected  $\tau_h$  efficiency VS jet fake-rate

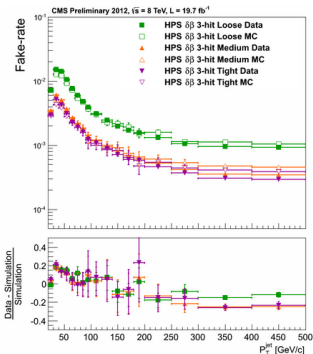
- MVA isolation reduces jet  $\rightarrow \tau_h$  fake rate by about a factor 2 compared to cut-based isolation
- The improvement is more relevant for  $\tau_h$  with high- $p_T$

Data/simulation comparison for  $\tau_h$  identification efficiency

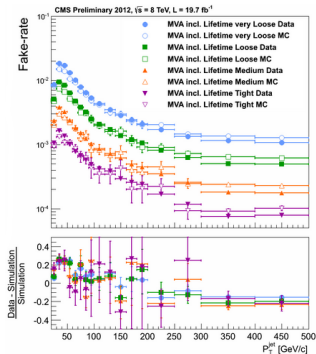
- $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$  events:  
1 muon ( $p_T > 25$  GeV); 1 jet ( $p_T > 20$  GeV,  $\Delta R > 0.5$ )  
opposite charge,  $m_T(\mu, \text{met}) < 40$  TeV, veto b-jet and extra leptons
- Events divided in 2 categories: pass or fail  $\tau_h$  identification
- Simultaneous fit in pass and fail categories for the track multiplicity, using  $\tau_h$  identification efficiency of data/simulation as parameter of interest
- About 70% efficiency for Loose WP of MVA discriminator
- $\tau_h$  efficiency of data/simulation is compatible with 1.0 within uncertainties ( $\sim 5\%$ )

Data/simulation comparison for jet  $\rightarrow \tau_h$  fake rate

## Cut-based isolation



## MVA-based isolation



- Multijet events:  
 $\geq 2$  jets
- Fake-rate measured in data and MC simulation as a function of  $p_T$
- Similar data/MC agreement for cut-based and MVA-based tau ID discriminators
- Fake rate varies from few per-cent (lower  $p_T$ ) to per-mille (higher  $p_T$ )
- Data/MC differences up to 20% observed. Some indication for a trend that the fake rates measured in data are higher (lower) compared to the MC simulation for low (high)  $p_T$  jets

# Tau identification against electrons and muons

- Electrons and muons may be reconstructed in the  $h^\pm$  decay mode or in the  $h^\pm \pi^0$  channel if electrons radiate a bremsstrahlung photon
- Define some  $\tau_h$  identification criteria against electrons and muons

## Electrons

- Energy amounts of ECAL and HCAL deposits
- Vars related to brems photons
- Track related variables:  $pT$ ,  $\eta$ ,  $\chi^2$ , num of hits in the tracked detector

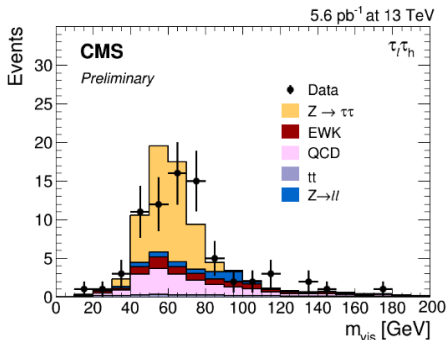
## Muons

- Energy amounts of ECAL and HCAL deposits
- $pT_\mu / pT_{\tau_h}$
- Num of hits and track segments in the muon chambers

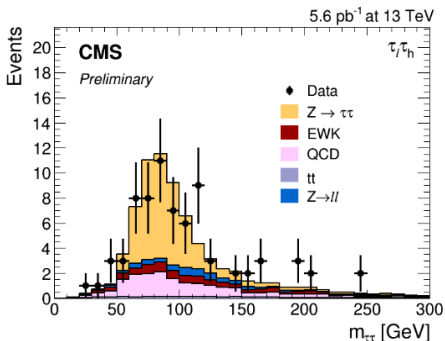
- The **rate** of electrons (muons) that are **mis-identified** as  $\tau_h$  is **measured** from data with a tag-and-probe technique in  $Z \rightarrow ee$  ( $Z \rightarrow \mu\mu$ ) events
- The measured  $e \rightarrow \tau_h$  ( $\mu \rightarrow \tau_h$ ) fake rate is **on the level of per-mille** for a  $\tau_h$  efficiency of **90%** (**98%**)

## First look at 13 TeV data in CMS

## Visible mass



## Fully reconstructed mass



- $Z \rightarrow \tau\tau \rightarrow e(\mu)\tau_h$  events:  
 $pT_e$  ( $pT_\mu$ ) > 20 (18) GeV and  $\eta < 2.1$   
 $pT_{\tau_h} > 20$  and  $\eta < 2.3$   
 Opposite charge and  $mT(\mu, \text{met}) < 40$  GeV
- Already **good performance in Run II**

# Conclusions

- HPS algorithm used in CMS provides powerful  $\tau_h$  reconstruction and identification
- Discriminators are introduced to reduce:
  - jet  $\rightarrow \tau_h$  fake rate, through tau isolation and lifetime variables
  - e  $\rightarrow \tau_h$  fake rate, through ECAL,HCAL energy information and track related variables
  - $\mu \rightarrow \tau_h$  fake rate, through ECAL,HCAL energy information and muon chamber related variables
- Good performances:
  - Tau identification between 50-60%
  - jet  $\rightarrow \tau_h$  fake rate about 1% (0.1%) for low pT (high pT) jets
  - Misidentification rate for electrons or muons on the order of per-mille
- Performance validated using  $19.7 \text{ fb}^{-1}$  of 8 TeV data recorded in 2012:
  - good data/MC agreement for  $\tau_h$  identification
  - up to 20% difference for jet mis-identified as  $\tau_h$
- Tau lepton was widely used during Run I allowing  $3 \sigma$  evidence for the SM Higgs in 2 taus  
(see Rainer Mankel's talk on Higgs results from CMS)  
SM measurements and beyond SM measurements  
(see Claudia Wulz's talk on Exotica in CMS)
- We had a first look at 13 TeV results,  
waiting for more data to perform complete measurements