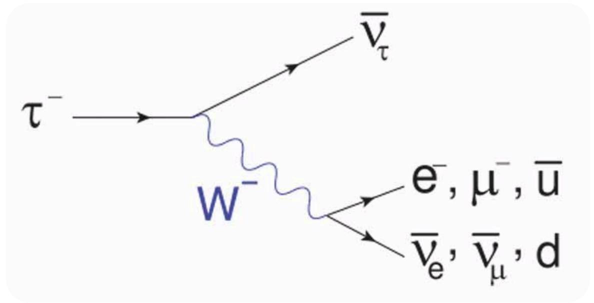
Tau Identification at CMS

Mauro Verzetti - University of Zurich
Winterthurerstr. 190 -8057 Zurich
on behalf of the CMS Collaboration

1. Tau lepton and its decay

The tau is the heaviest (known) lepton with a mass of 1.78 GeV. Its decay is electroweak (thus involving **neutrinos**) with a cτ of 87 μm.

Tau decays into other leptons about 17% of the times, while the rest of the decays is hadronic (in future called τ_h), mainly involving π's.

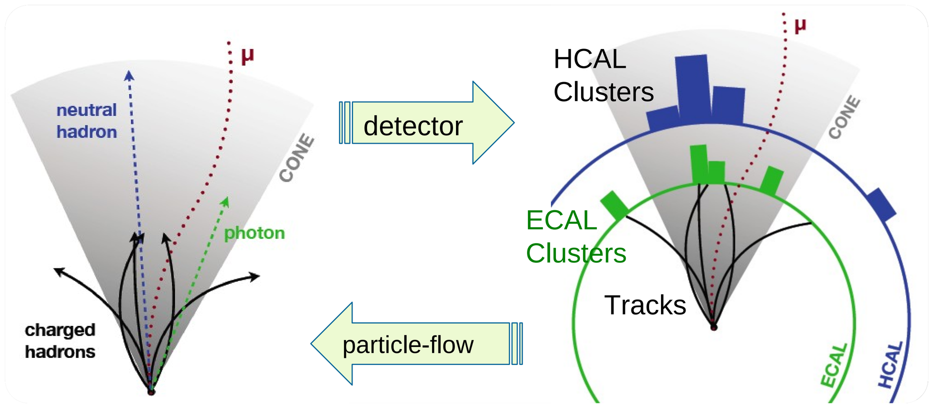


Hadronic tau decays are usually divided according to the number of charged particles that they involve (one or three). Due to the low number of decay products and the large boost of taus at LHC, tau jets can be identified exploiting low detector activity around decay products (aka **isolation**)

2. Particle Flow

Particle Flow [1] is an algorithm that gives a complete description of the event.

Linking all the signals coming from different subdetectors and the output of the reconstruction algorithm of CMS, Particle Flow produces a list of particle candidate (e, μ, γ, neutral and charged hadrons) that look very similar to a simulation for the end user.

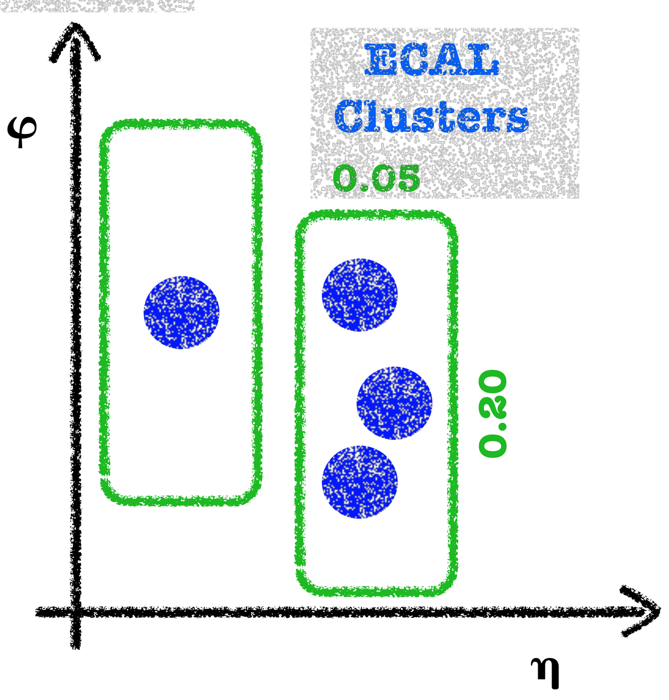


Hadronic tau identification is built inside this framework in order to exploit the best knowledge of the event available.

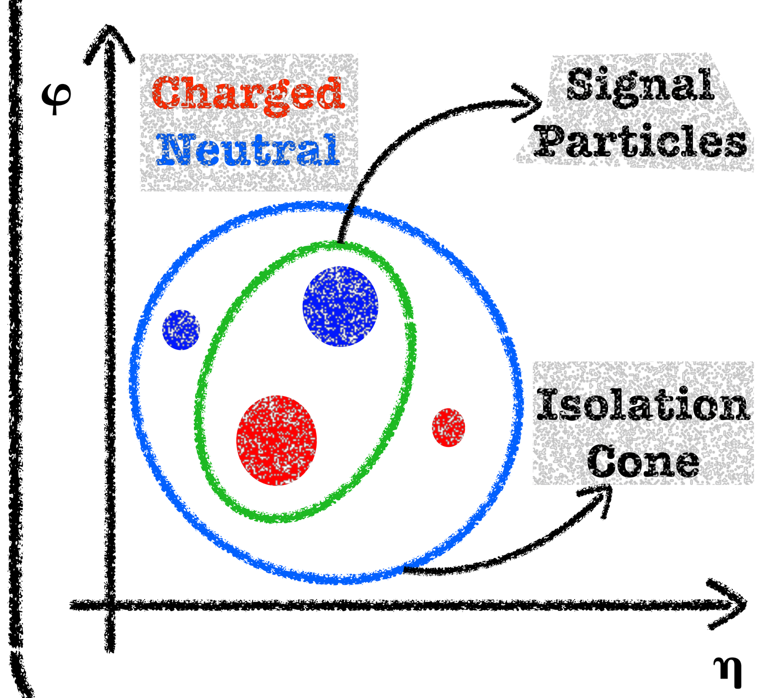
3. Hadron Plus Strip

The main algorithm to reconstruct taus at CMS is Hadron Plus Strip (**HPS**) [2].

Starting from a jet the algorithm initially clusters the ECAL clusters opening a strip in φ. This way it is possible to recover the energy of photons that convert in the detector. Reconstructed photons are then used to build π^0 candidates.



The hadron and photon candidates inside the jet are then combined to form a candidate τ decay, requiring also that the system is compatible with a ρ or a_1 hypothesis.



The most isolated combination is then chosen as tau candidate for that jet.

Isolation is computed looking at the remaining candidates inside the jet and either counting the number of those which are over a certain p_t threshold or simply summing their p_t. **Three different working points are provided with different isolation requirements.**

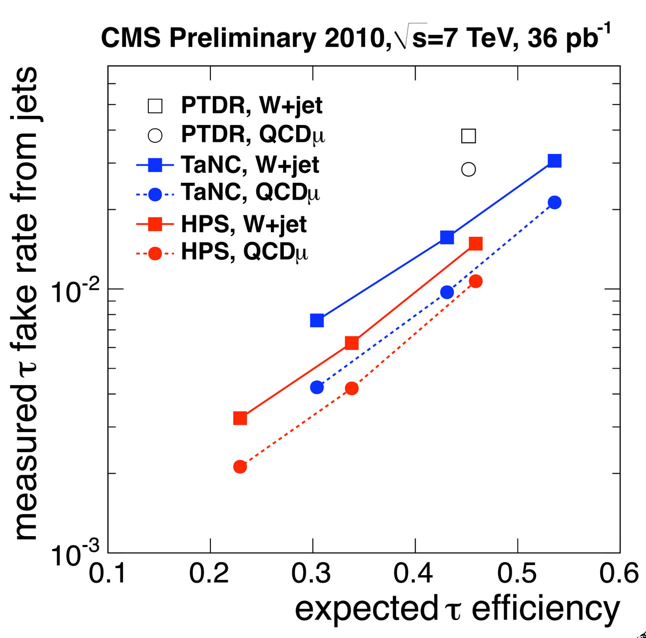
A final cross cleaning step against muons and electrons is also applied.

4. Fake rates

The rate of jets faking hadronic tau decays has been measured in samples of QCD and W+jets events.

The plot on the right shows the fake-rates measured in CMS data recorded in 2010 for Loose, Medium and Tight working-points of the HPS algorithm (red), compared to alternative CMS tau identification algorithms (black, blue).

On the x-axis efficiencies predicted by the Monte Carlo simulation are shown.

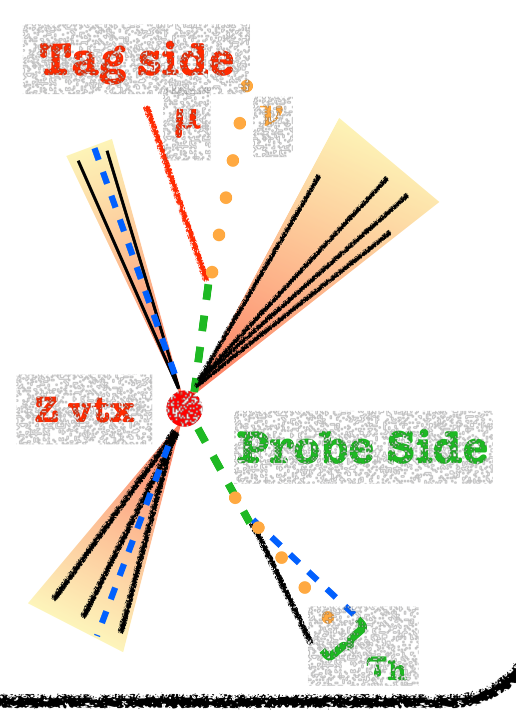


5. Measuring tau identification efficiency

The aim of this analysis is to measure the algorithm efficiency from data in the most unbiased way.

Three methods are used:

- Ratio of Z→ττ over Z→μμ yield
- Measuring Z→ττ→μ(e)+Jet and Z→ττ→μ+μ(e)
- **Tag and probe method (presented in this poster)**



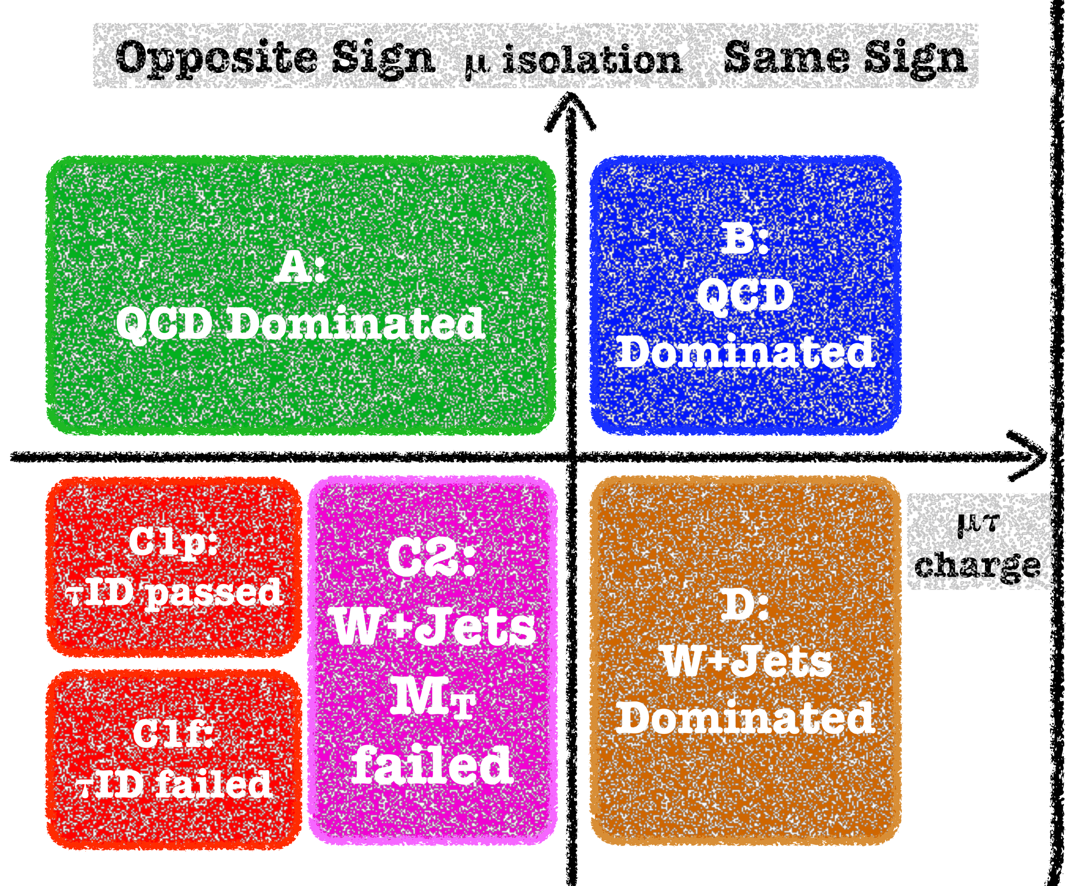
Tag and probe is the only one providing a measurement even for analyses searching for ττ excess in the Z peak region (like Higgs searches [3]).

6. Event selection

All the events are required to satisfy the following requirements:

- p_T^{jet} > 20 GeV, |η| < 2.3
- Jet leading track p_T > 5 GeV
- Loose isolation for muon and Jet candidates

Events are then divided into four different regions according to the muon isolation, the sign of muon and leading track of the jet. The signal region is then additionally purified requiring that the transverse mass of the muon plus the MET is less than 40 GeV (M_T in the cartoon), increasing the number of regions to five.



Finally the **Tau ID is applied** dividing the region in **passed** and **failed**.

7. Fitting Model

Several sources of background were considered: W+Jets, Z→μμ, QCD μ enriched (each event contains a muon) and one source of signal: Z→ττ.

Monte Carlo simulations were used to determine the lineshapes of the different processes in each region. Only the shape of **QCD** in **C1** regions was taken from data.

Distributions of μ + τ visible mass (in regions **C1p**, **C1f**) and of the transverse mass of μ + MET (in regions **A**, **B**, **C2**, **D**), were computed for each source and the data.

Each region was then fitted using the MC shapes as template, constraining the normalizations to be consistent among all regions (**ABCD** method).

Efficiency is then computed as:

$$\varepsilon = \frac{N^{fit}_{C1p\ Z\rightarrow\tau\tau}}{N^{fit}_{C1p\ Z\rightarrow\tau\tau} + N^{fit}_{C1f\ Z\rightarrow\tau\tau}}$$

8. Systematics

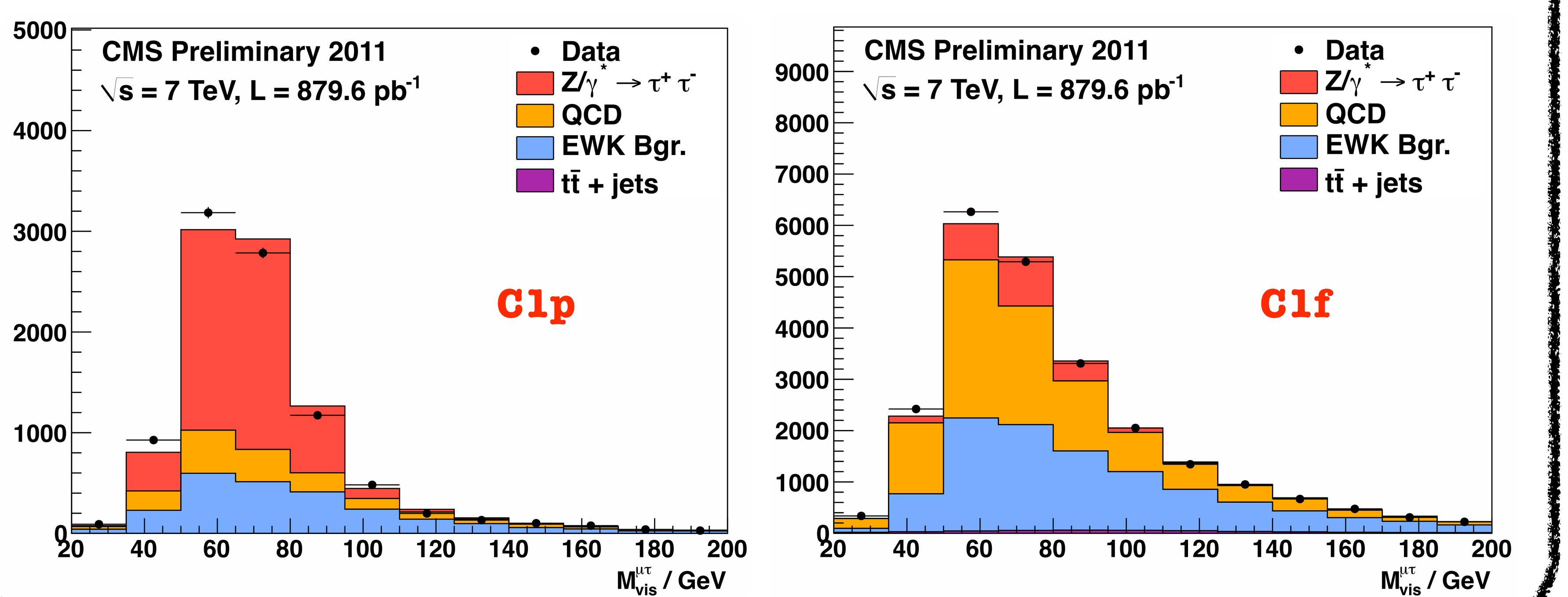
Sources of systematic uncertainties are:

- Hadron track reconstruction efficiency (3.9%)
- Uncertainty on the correction factor due to contamination of jets faking taus in the Z→ττ Monte Carlo template (1.2%)
- Uncertainty on the preselection cut efficiencies (1.6% for leading track, 2.1-1.5% for loose isolation of the jet)

9. Results

The measured tau identification efficiency was found to be in agreement with the Monte Carlo predictions.

Including statistical and systematic uncertainties, the precision of the measurement is 6%.



References:

[1] The CMS Collaboration: "Particle Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET" [CMS PAS PFT-09-001, http://cms-physics.web.cern.ch/cms-physics/public/PFT-09-001-pas.pdf].

[2] The CMS Collaboration: "Tau identification in CMS" [CMS PAS TAU-11-001, http://cms-physics.web.cern.ch/cms-physics/public/TAU-11-001-pas.pdf].

[3] The CMS Collaboration: "H ->Tau Tau" [CMS PAS HIG-11-009].