Tau Identification with Deep Neural Networks at the CMS Experiment

Somnath Choudhury

(for the CMS collaboration)

IEEE – RT 2020: 22nd IEEE Real Time Conference, 12 – 23 October 2020, Quy Nhon (Vietnam) – online mode

Tau Leptons

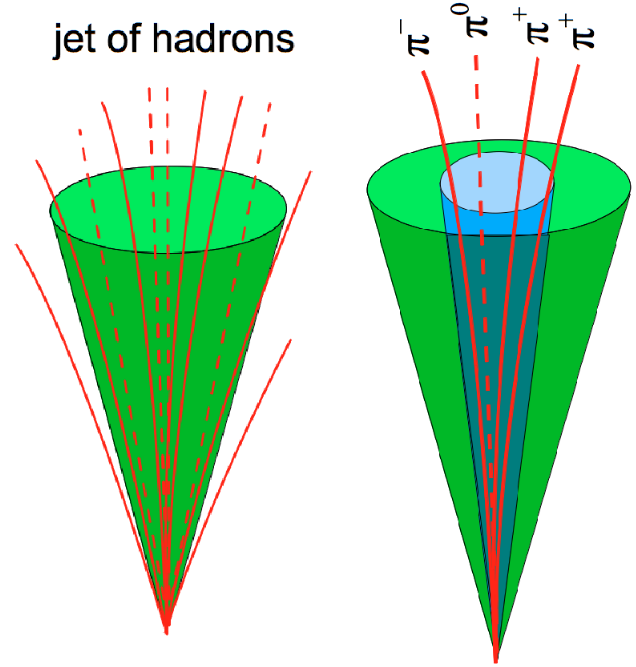
Tau plays crucial role in precision tests of electroweak physics and lepton flavour universality, studies of the 125 GeV Higgs boson via its Yukawa coupling to fermions, searches for additional Higgs bosons, several searches for BSM signatures like supersymmetry, dark matter and exotic heavy particles.

mass 1.776 GeV
⇒ can decay to hadrons (m_π = 140 MeV)

decay:
65% hadronic ("τ_h")
35% fully leptonic

lifetime τ = 2.9×10^{-13} s
⇒ $\gamma c\tau$ ~ **1 mm** (20 GeV)
⇒ secondary vertex

τ_h is a collimated, **isolated jet**



decay ~11.5% of the time to one charged hadron, 35.5% to one charged hadron and neutral pions, 15% to three charged hadrons and neutral pions.

- 1st step: Hadronic tau candidates reconstructed via the Hadron Plus Strips algorithm **JINST 13 (2018) P10005** ⇒ searches for neutral pions and charged particles within the jet.

- 2nd step: The tau candidate is then accepted into one of four categories, depending on the number of neutral pions and charged particles found.

- 3rd step: Feeding the tau candidates into a deep neural network based algorithm to remove fake taus.

- discriminators make mistakes and reject true taus, or accept fake taus as true ⇒ balancing between efficiency (true acceptance) and misidentification rate is done by defining **working points**.

- **working points** from a loose selection (large number of accepted taus, with large amount of fake taus), to a tight selection (higher purity of true taus, but smaller sample, as true taus are cut out too).

DeepTau Algorithm

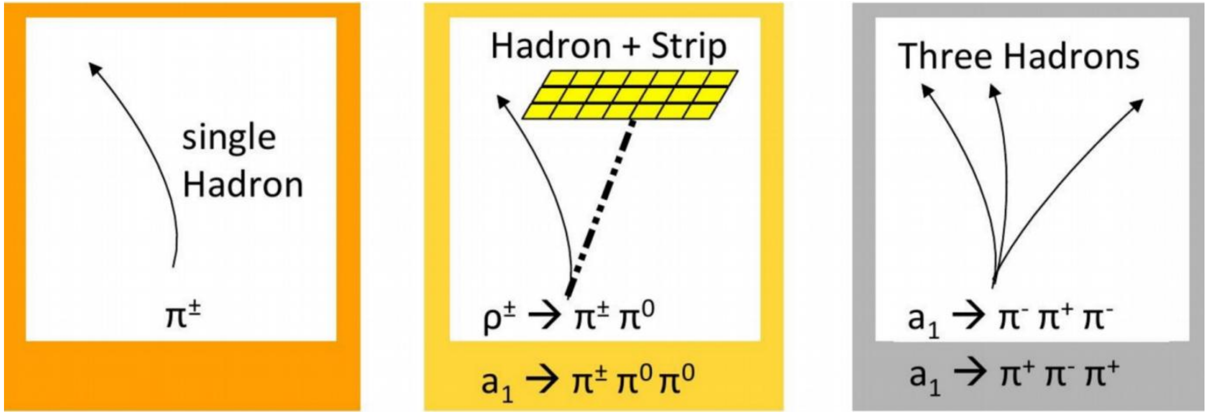
DeepTau => new multiclass tau identification algorithm based on a convolutional deep neural network (DNN)

The DeepTau convolutional neural network (NN) used to **reduce the misidentification** of quark / gluon jets, muons and electrons as hadronic taus

DeepTau combines information from high-level reconstructed tau features together with low level information from the inner tracker, calorimeters and muon sub-detectors using particle flow (PF) candidates, electrons and muons reconstructed within the tau isolation cone

Inputs:

- **Low-level**
  Tracks and energy deposits of PF candidates
- **High-level**
  Transverse momenta, decay mode, etc. of tau candidate + general event properties



DeepTau Algorithm

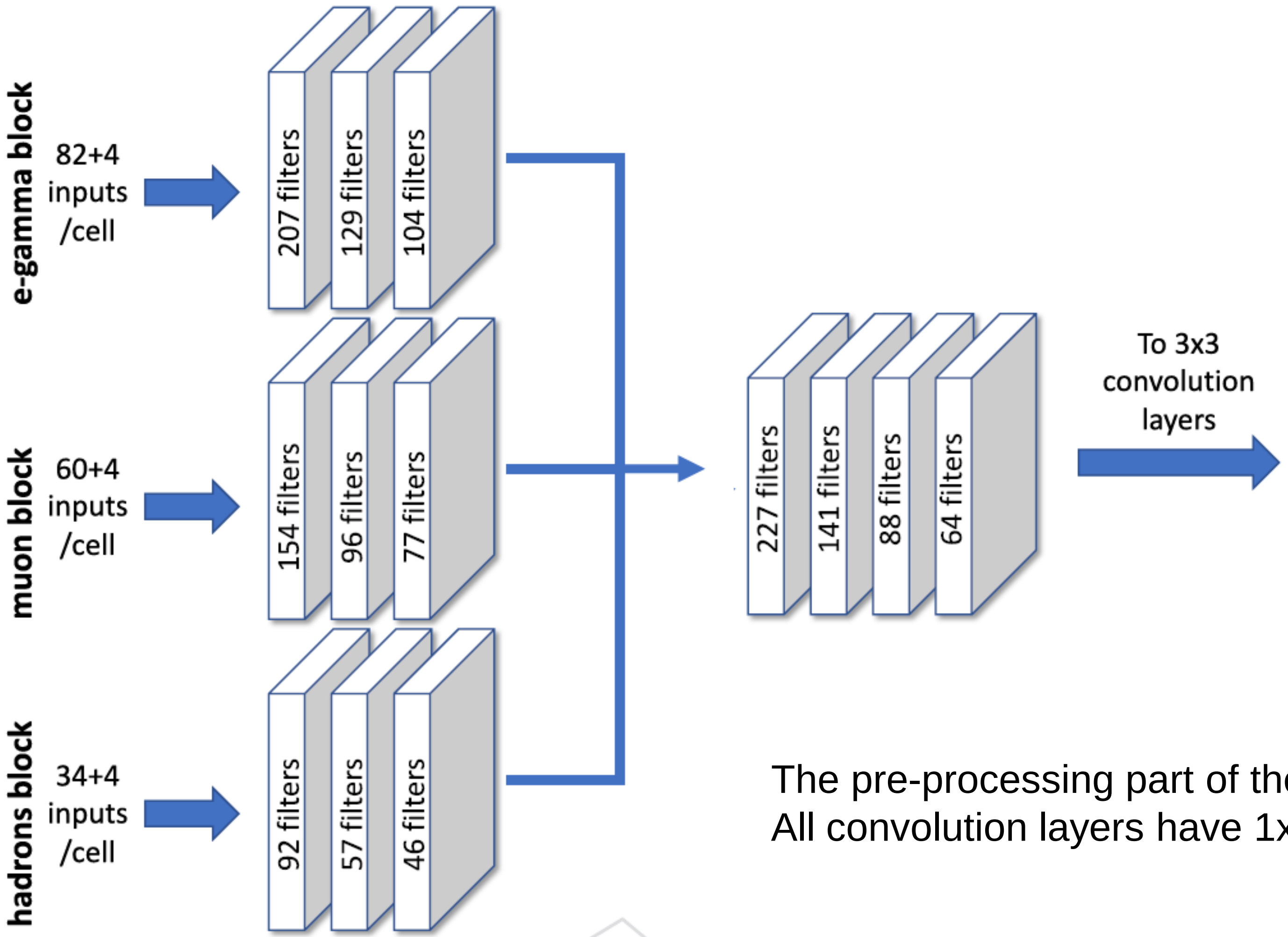
The organisation of low-level inputs into two 2D grids:

- first process the local patterns originating from the tau or jet structure
- then iteratively combine obtained information covering bigger $\eta \times \varphi$ regions up to the point where the whole tau signal or isolation cones are covered

Approach inspired by similar techniques used in modern Machine Learning based image recognition with convolutional DNNs

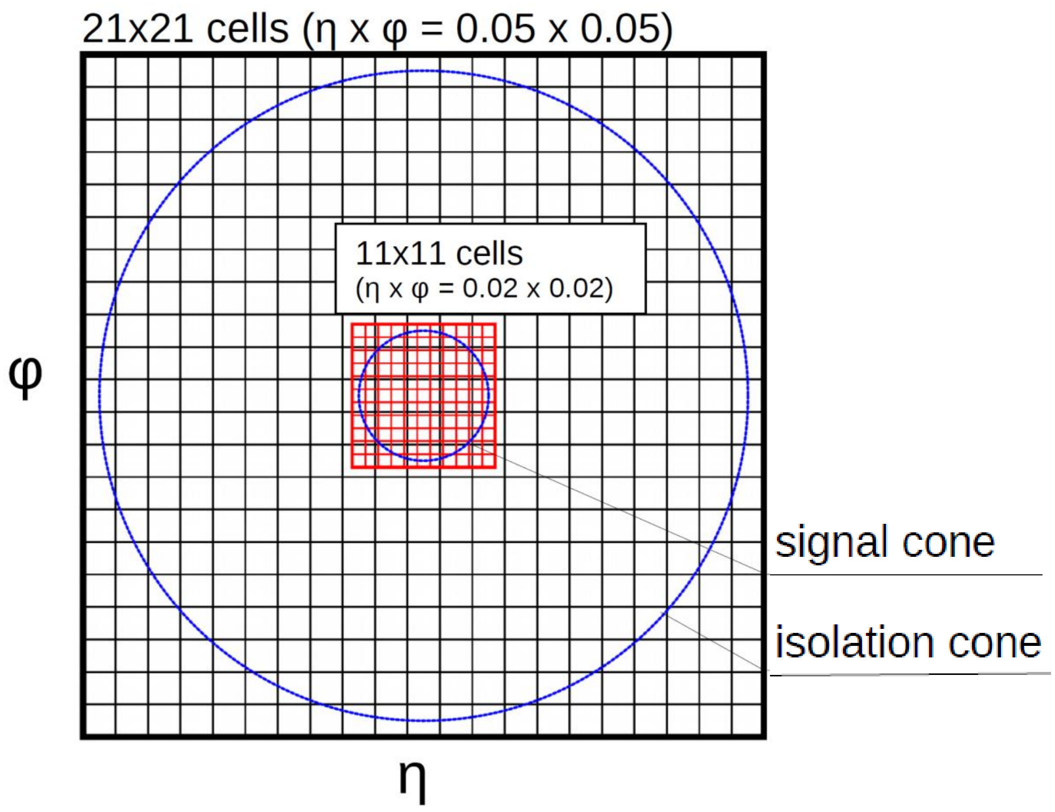
- **Pre-processing:** The outputs of all 3 blocks of electrons-photons, muons and hadrons are concatenated and pre-processed together by 4 convolutional layer.
- **Convolution:** As the output of the pre-processing step, we have 11x11x 64 array for the inner grid and 21x 21x 64 array for the outer grid. These arrays processed by several convolution layers with 3x3 window size.
- **Final merge:** Outputs from tau-pre, inner and outer convolution blocks concatenated (producing 185 inputs) and processed by 5 connected dense layers. Number of units in each layer is fixed to 200.
- **Outputs:** Final dense layers transforms 200 features into 4 outputs of network → represent estimates of the probabilities of the tau candidate to be an **electron**, **muon**, **jet** or **genuine hadronic tau**.

- High level variables and each input cell are pre-processed by a few fully connected dense layers



The pre-processing part of the DeepTau NN.
All convolution layers have 1x1 window size.

- For the inner (outer) grid, the pre-processed cell data are fed into 5 (10) 2D convolutional layers with 3×3 window size, which result in 64 features that are passed to the next step



● Low level inputs are based on the tau decay products

● CMS detector divided in cells of pseudorapidity η x azimuthal angle φ

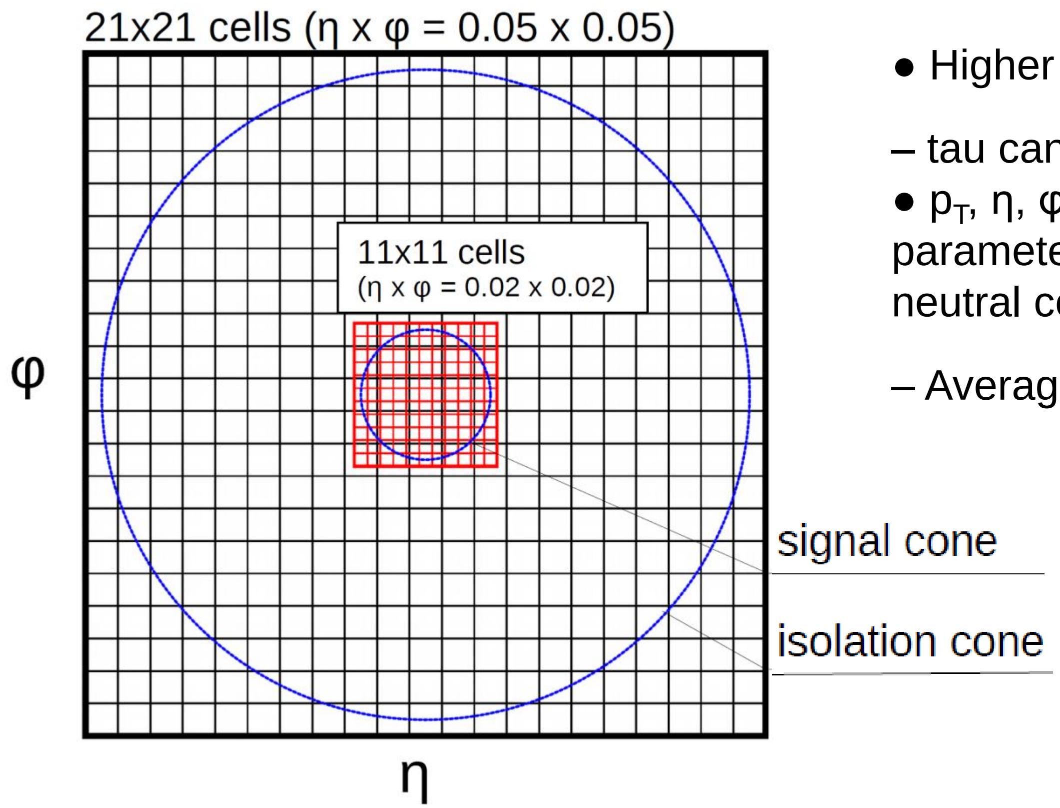
● In each cell all available information for the leading PF candidate stored

Candidates belonging to the inner and outer cones are separated and split into two grids with **η x φ** cell size of

- ✓ **0.02×0.02** (**0.05×0.05**) for the inner (outer) cone

- Higher density cells in the **signal cone** ($\Delta R < 0.1$)

- Coarser set of cells in a $\Delta R < 0.5$ cone → **Isolation cone**



- 2 different granularities chosen to reduce the number of inputs for the NN (188 features)

- Higher level inputs (47 features):

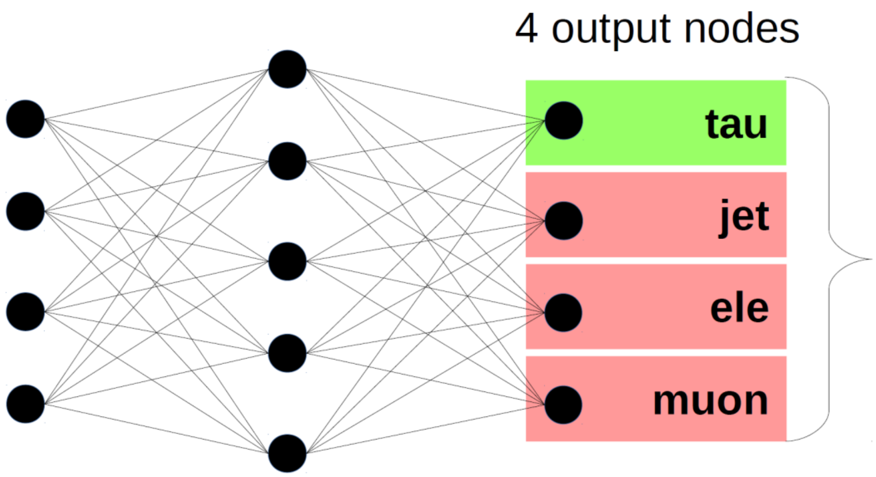
– tau candidate properties:

- p_T, η, φ, decay mode, charge, impact parameter, number of charged prongs and neutral constituents, etc.

– Average energy in the event, Δη ECAL, etc.

Network Architecture

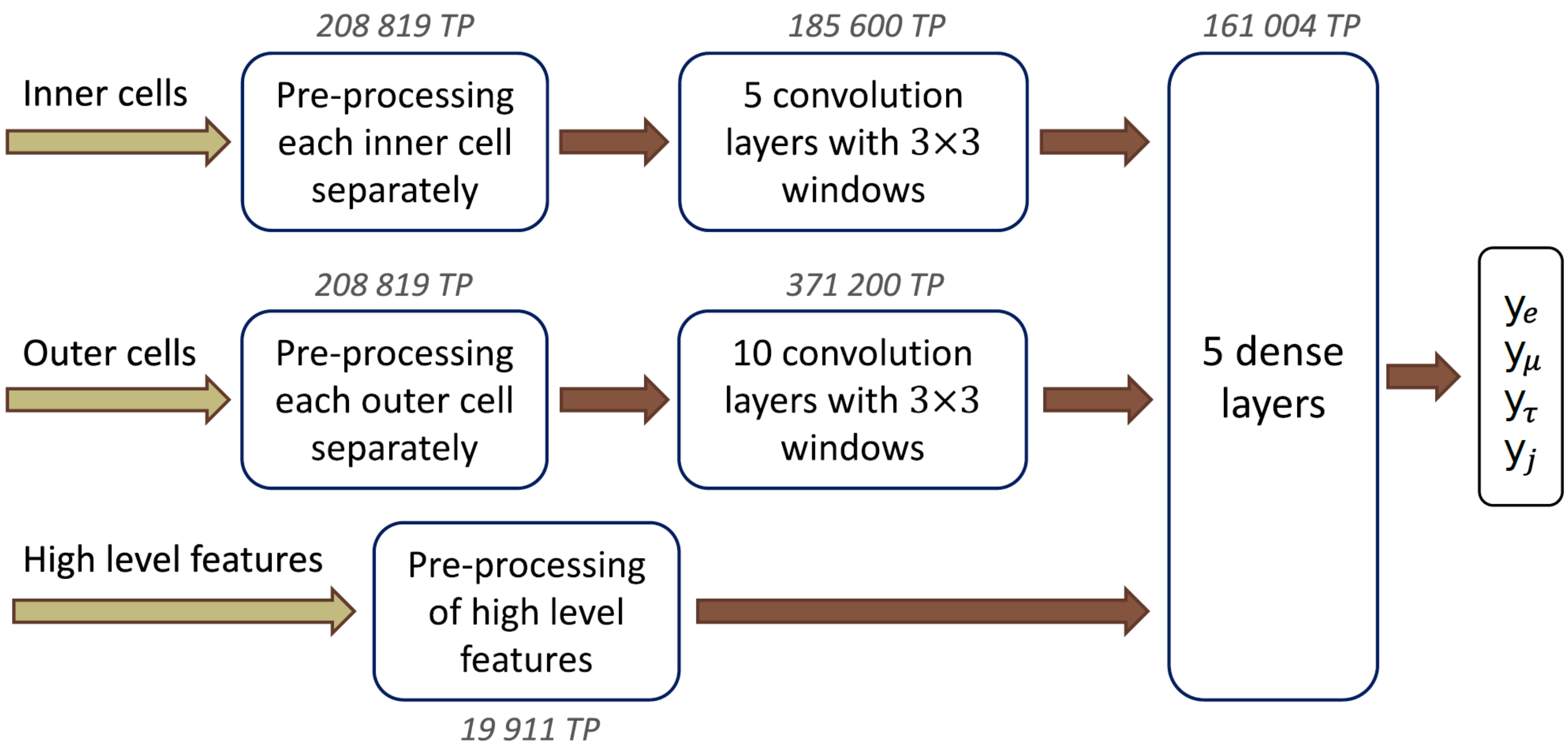
- All features from the previous steps are combined and passed through 5 dense layers
- Probabilities of the reconstructed tau candidate being electron, muon, quark or gluon jet, or genuine hadronic tau are estimated by the 4 NN outputs



$$P_{\tau\, vs\, obj.} = \frac{P_{\tau}}{P_{\tau} + P_{obj.}}$$

Classifiers used to reduce the misidentification rate

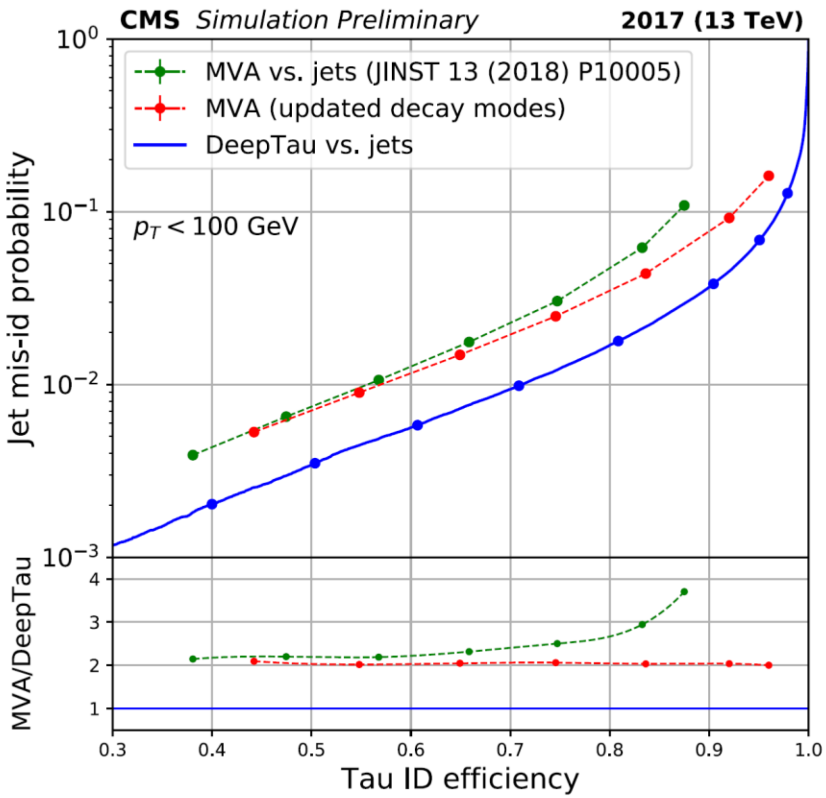
The final NN architecture chosen for DeepTau algorithm. The number of trainable parameters (TP) are indicated for each group of layers.

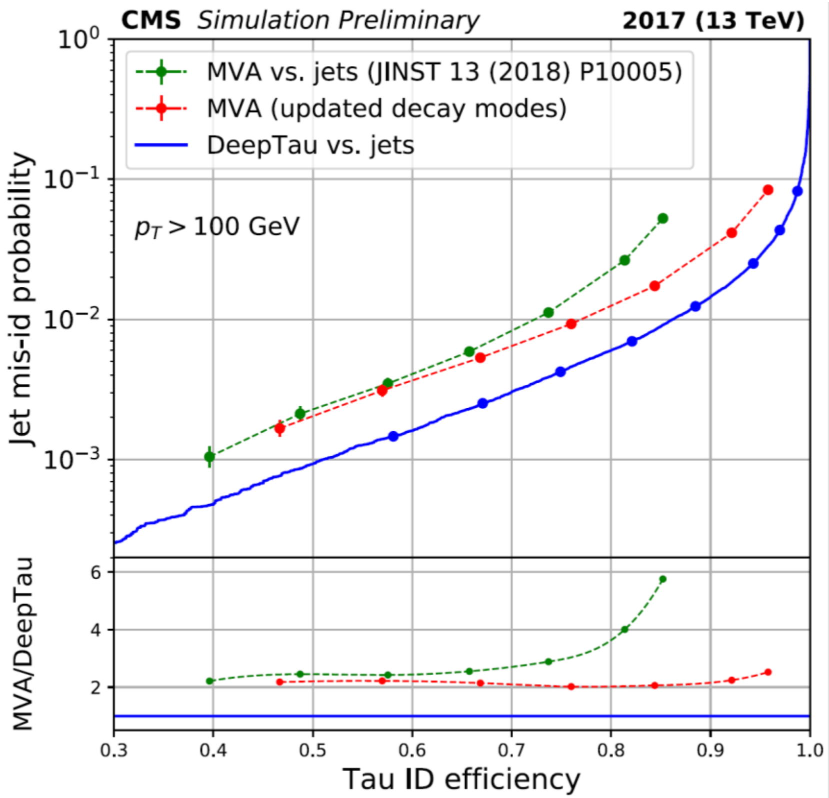


Discrimination against Jets

DeepTau performance evaluated using MC simulation under 2017 collision conditions on reconstructed τ candidates in p_T range 20-1000 GeV
Tau ID efficiency estimated from **H→ττ** MC using reconstructed τ candidates matching hadronically decaying τ at generator level.
Jet misidentification probability estimated from **top pair** MC using reconstructed τ candidates not matching hadronically decaying τ at generator level.

CMS DP -2019/033





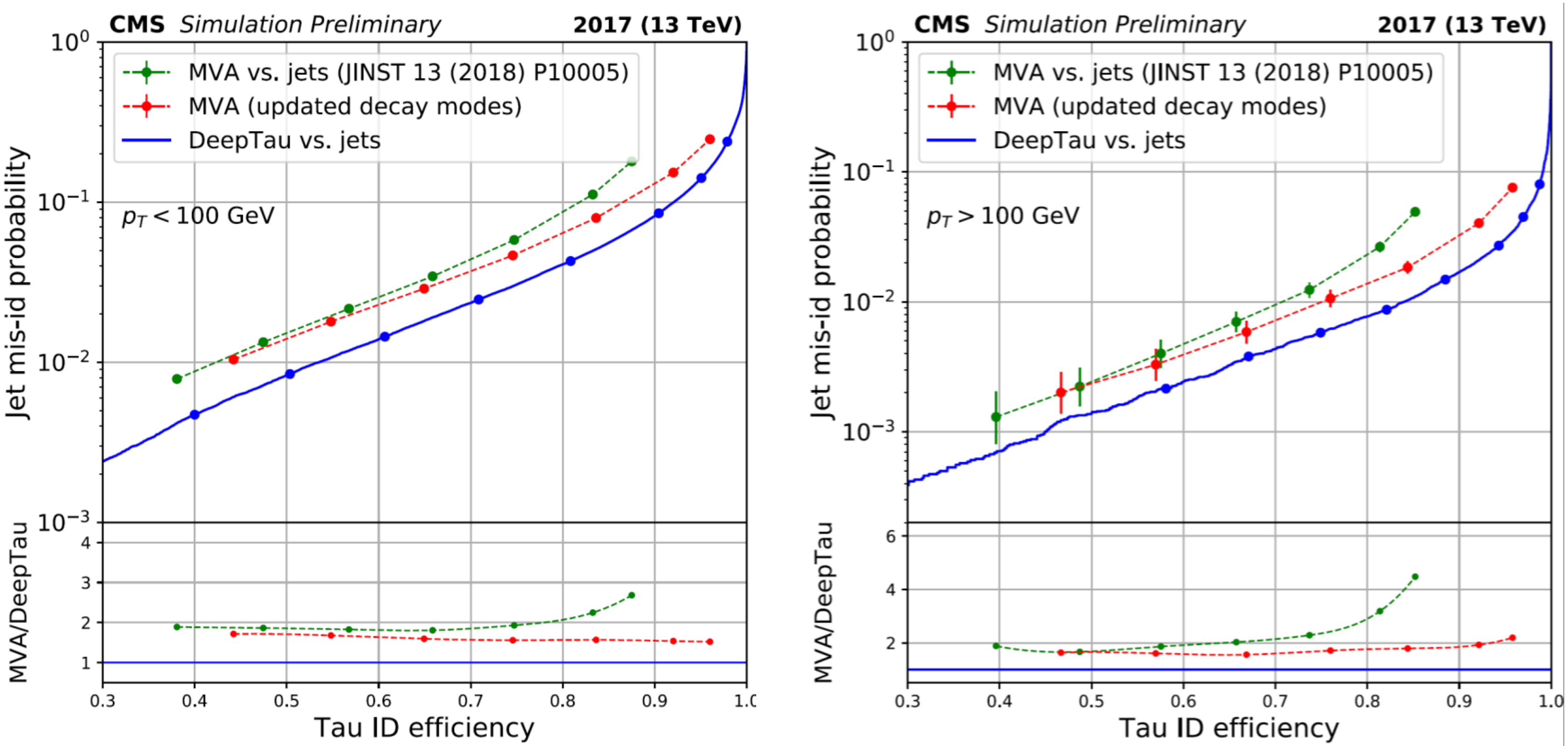
Discrimination against Jets

DeepTau performance evaluated using MC simulation under 2017 collision conditions on reconstructed τ candidates in p_T range 20-1000 GeV

Tau ID efficiency estimated from **$H\rightarrow\tau\tau$** MC using reconstructed τ candidates matching hadronically decaying τ at generator level.

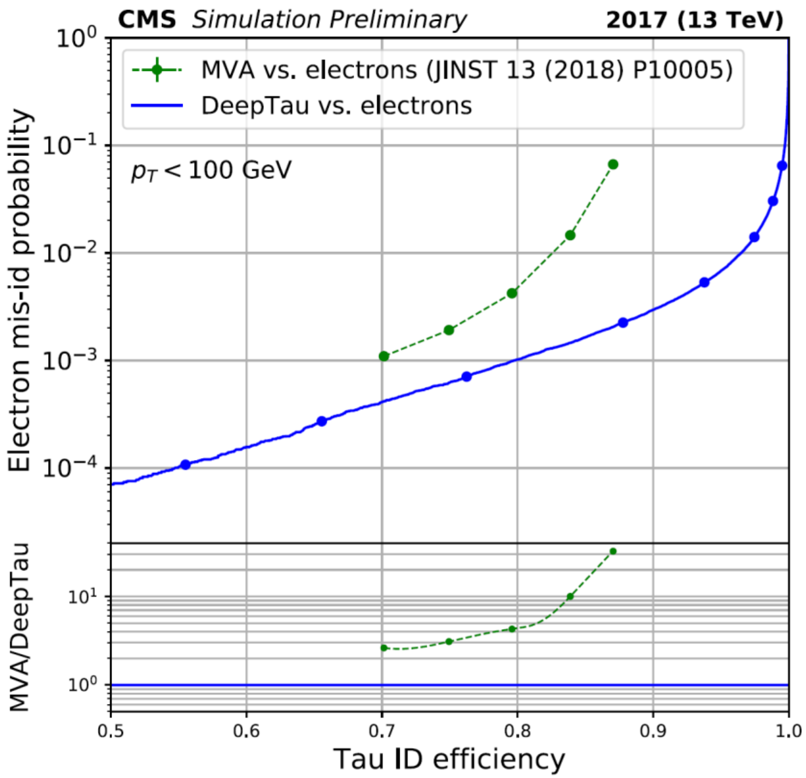
Jet misidentification probability estimated from **W+jets** MC using reconstructed τ candidates not matching hadronically decaying τ at generator level.

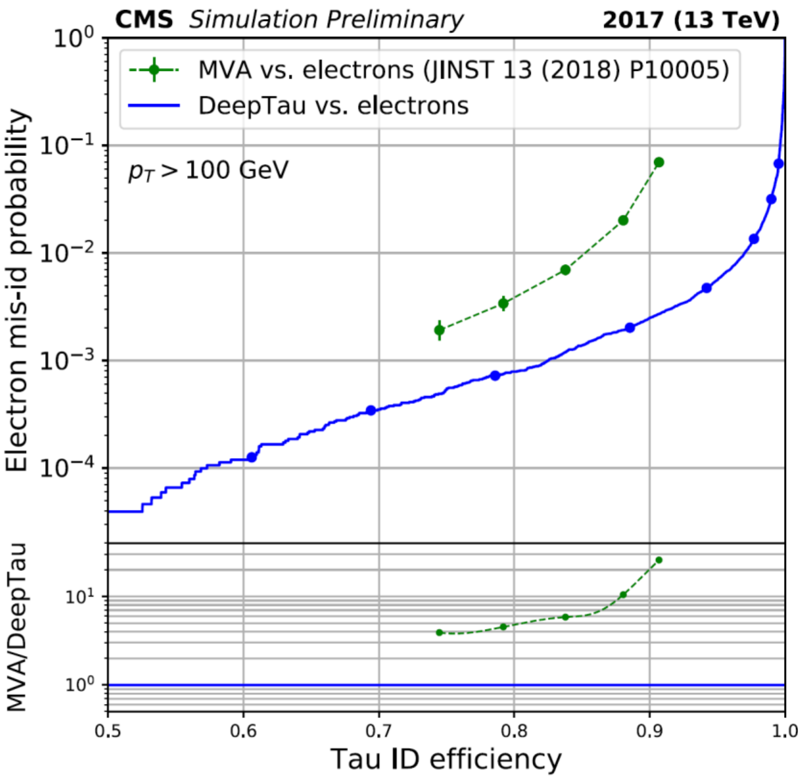
CMS DP -2019/033



Electrons which can emit photons via bremsstrahlung radiation and mimic tau decay via the ρ resonance.
DeepTau performance evaluated using MC simulation under 2017 collision conditions.
Tau ID efficiency estimated from **H→ττ** MC. Electron misidentification probability is estimated from Drell-Yan **Z/γ*→ee** MC using reconstructed tau candidates that match electrons at the generator level.

CMS DP -2019/033



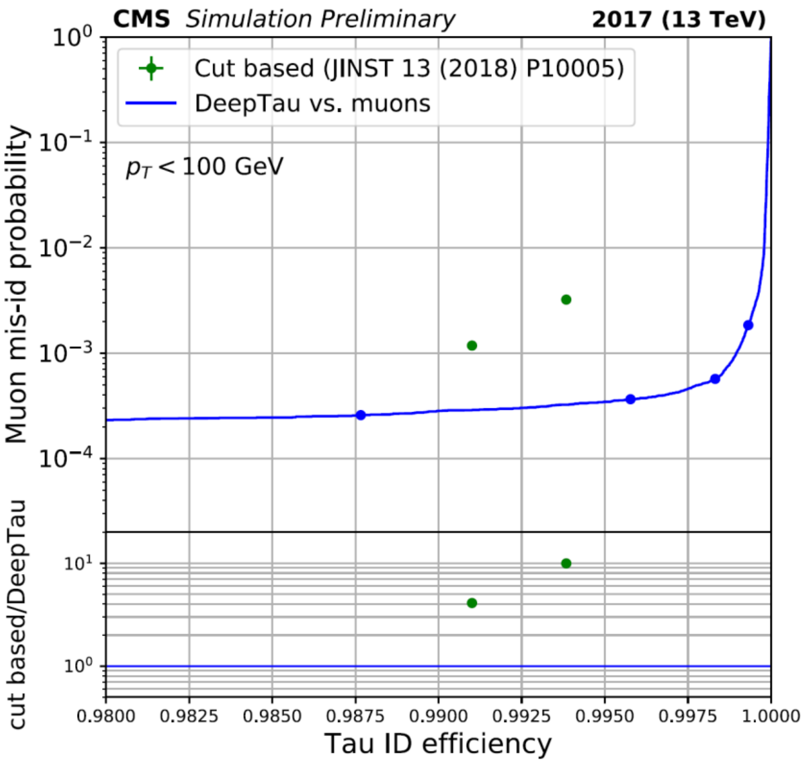


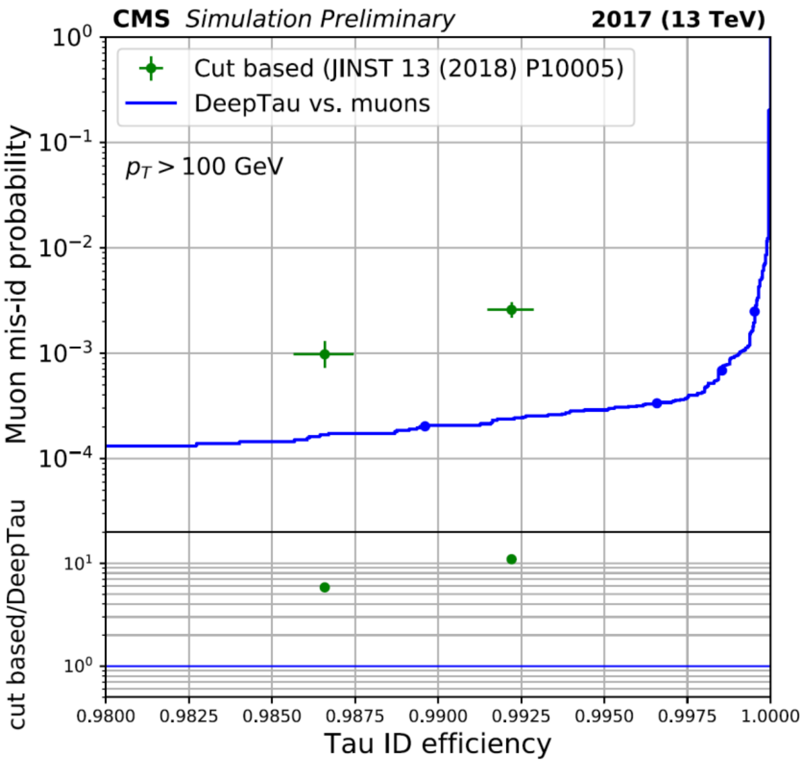
Muons can be misidentified as tau by producing a signature similar to tau decaying to a charged hadron.

DeepTau performance evaluated using MC simulation under 2017 collision conditions.

Tau ID efficiency estimated from **$H \to \tau\tau$** MC. Muon misidentification probability is estimated from Drell-Yan $Z/\gamma^* \to \mu\mu$ MC using reconstructed tau candidates that match muons at the generator level.

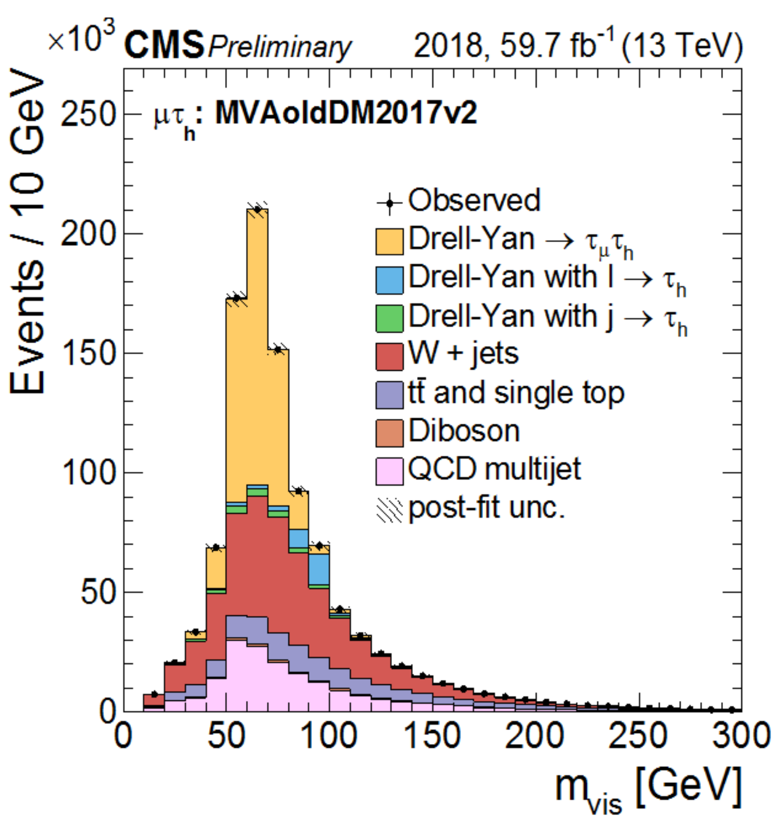
CMS DP -2019/033

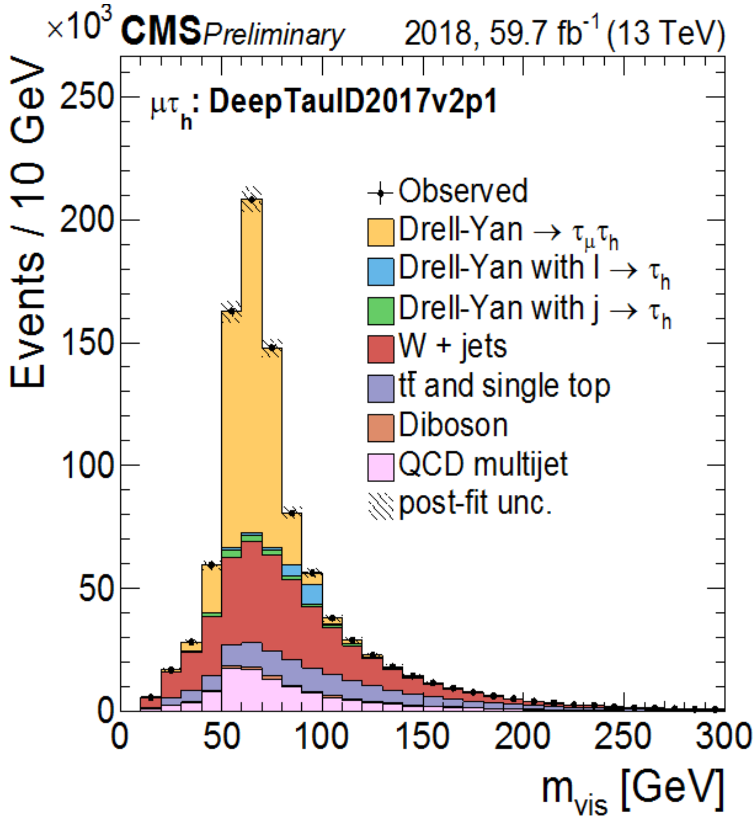




Distribution of the visible tau pair mass in semileptonic decay to muon and hadrons is studied with DeepTau algorithm and the previously used MVA algorithm for 2018 proton collision data. Contribution from all SM processes (except QCD) modelled by MC. QCD estimated from a sideband region in data.

With DeepTau yield from **genuine τ_h** increases by ≈**20%**, while yield from **fakes** decreases by ≈**23%**





Summary

Tau leptons very important for new physics searches and precision measurements as tests of SM.

Excellent performance and understanding of the CMS detector allows for a good description of τ_{h} reconstruction in spite of the dense hadron environment of the LHC.
Particle Flow technique at CMS provides well-calibrated constituents of τ_{h} decay

A new deep neural network based algorithm **DeepTau** developed at CMS. The algorithm outperforms the previously used tau ID algorithms in CMS.

DeepTau algorithm tested with LHC Run-2 proton collision data shows **excellent performance** for discrimination against jets, electrons and muons.
With DeepTau ID, a **significant reduction in the misidentification rate** observed in data with the same identification efficiency compared to the previous algorithms.

DeepTau constitutes an extremely useful tool in reducing the background in several physics studies at CMS involving tau final states.