Identification of hadronic tau decays using a deep neural network with the CMS experiment at LHC

International Conference of High Energy Physics

ICHEP 2022

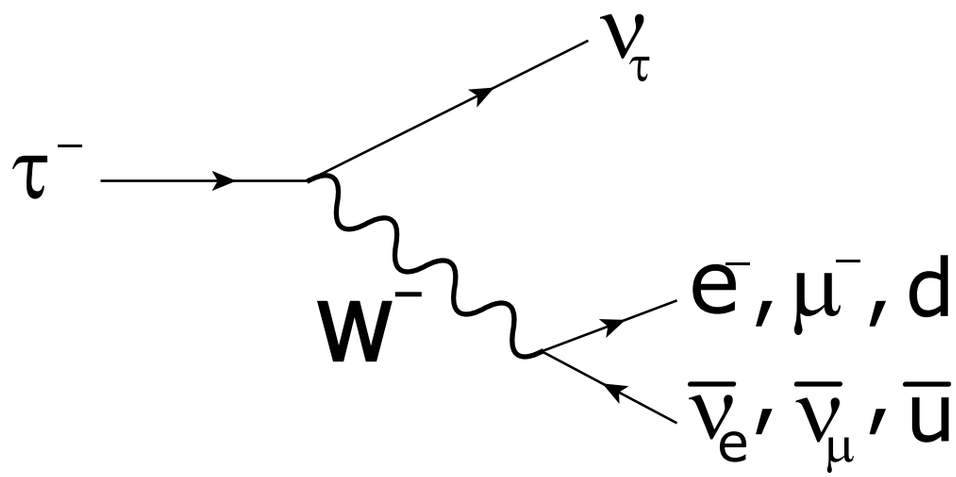
July 8th, 2022

Konstantin Androsov[1] on behalf of the CMS Collaboration

[1]EPFL and ETHZ

Analysis with hadronic taus at CMS

- **Tau** is the heaviest Standard Model (SM) lepton with the mass of 1.78 GeV that **decays into hadrons + neutrino in 64.8% of the cases**



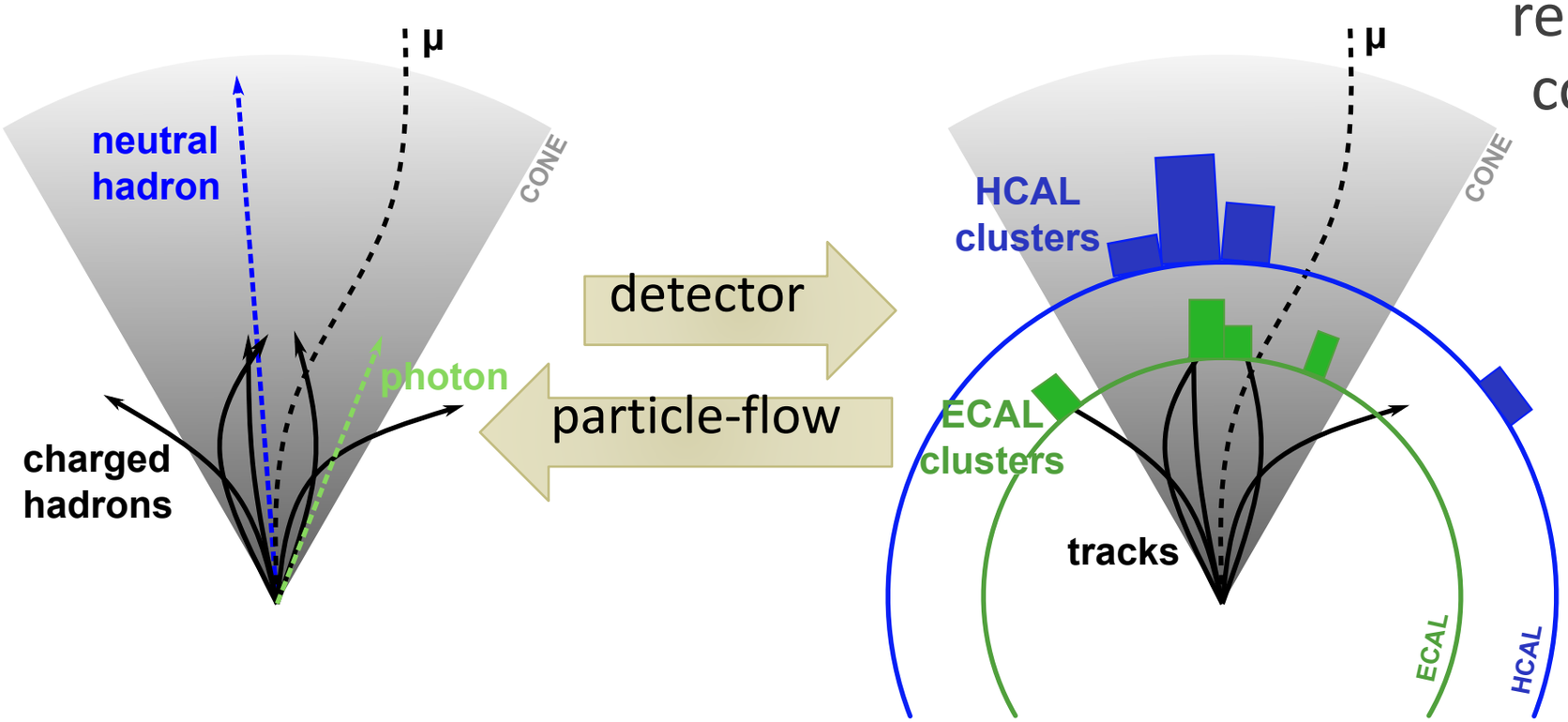
- Good performance in reconstruction and identification of the hadronic tau decays is crucial for many important physics analysis in CMS:
  - Measurement of the SM properties, including $H \to \tau\tau$ analyses with different production mechanisms
  - SM and BSM searches for Higgs pair production: $HH \to bb\tau\tau$, $HH \to 4\tau$, ...
  - Searches for leptoquarks, heavy neutral leptons, additional charged and neutral scalar bosons
  - *And many other CMS analyses*

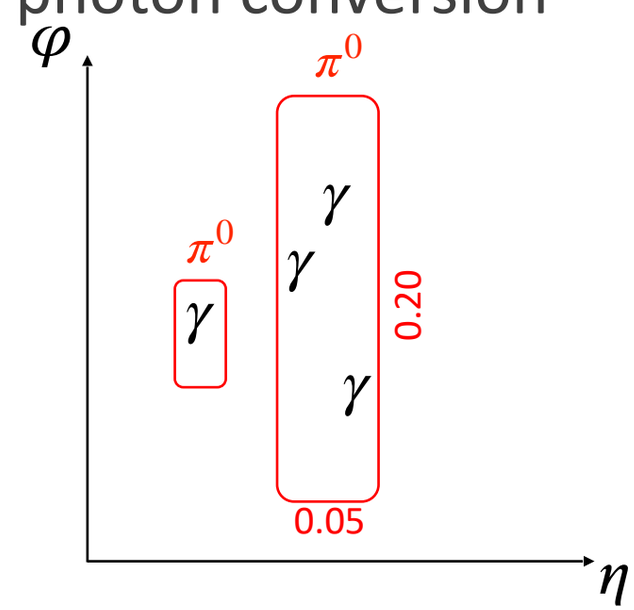
Reconstruction of the hadronic tau decays

- In CMS, stable particles are reconstructed using the **particle flow** (PF) algorithm
- Hadronic tau decays are reconstructed by **Hadron + strips** (HPS) algorithm
- Starting from jet, the HPS algorithm analyses information from stable PF particles within the jet cone to identify typical hadronic tau decay signatures

Decay mode	Resonance	$\mathcal{B}$ (%)	
Leptonic decays		35.2	
$\tau^- \to e^- \overline{\nu}_e \nu_\tau$			17.8
$\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$			17.4
Hadronic decays		64.8	
$\tau^- \to h^- \nu_\tau$			11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$		25.9
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$		9.5
$\tau^- \to h^- h^+ h^- \nu_\tau$	$a_1(1260)$		9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$			4.8
Other			3.3

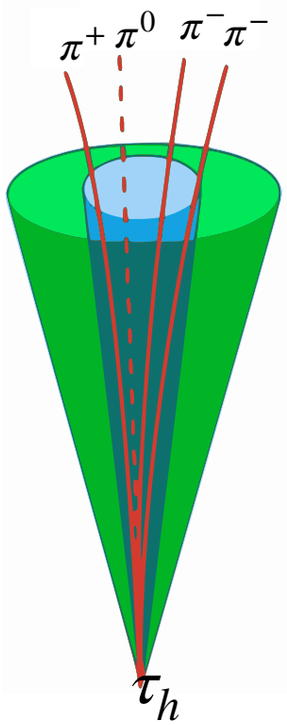
In HPS, neutral pions are reconstructed using $\eta \times \varphi$ strips, collecting energy spread from photon conversion

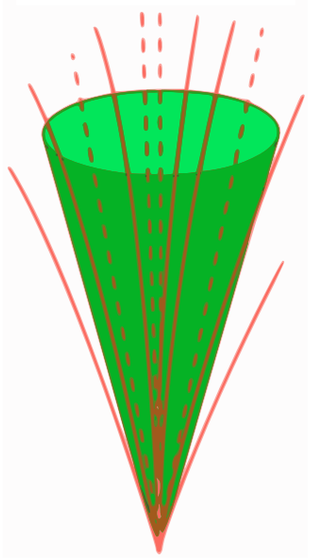




Main backgrounds for hadronic taus

- **Jets** originating from quarks or gluons (τ_j), **electrons** (τ_e), and **muons** (τ_μ) can be misidentified as **hadronic tau** decays (τ_h)
- Each background have a characteristic signatures in the detector that can help to separate it from τ_h:
  - τ_j candidates, in general, have more hadronic activity, which can be detected in the isolation cone
  - τ_e candidates have specific patters in the calorimeter clusters
  - τ_μ candidates have substantial amount of matched hits in the muon chambers
- For each case, based on these typical signatures a set of the most discriminating variables can be defined
- Before DeepTau, 3 dedicated algorithms were developed within CMS to discriminate τ_h against given source of the background (*)



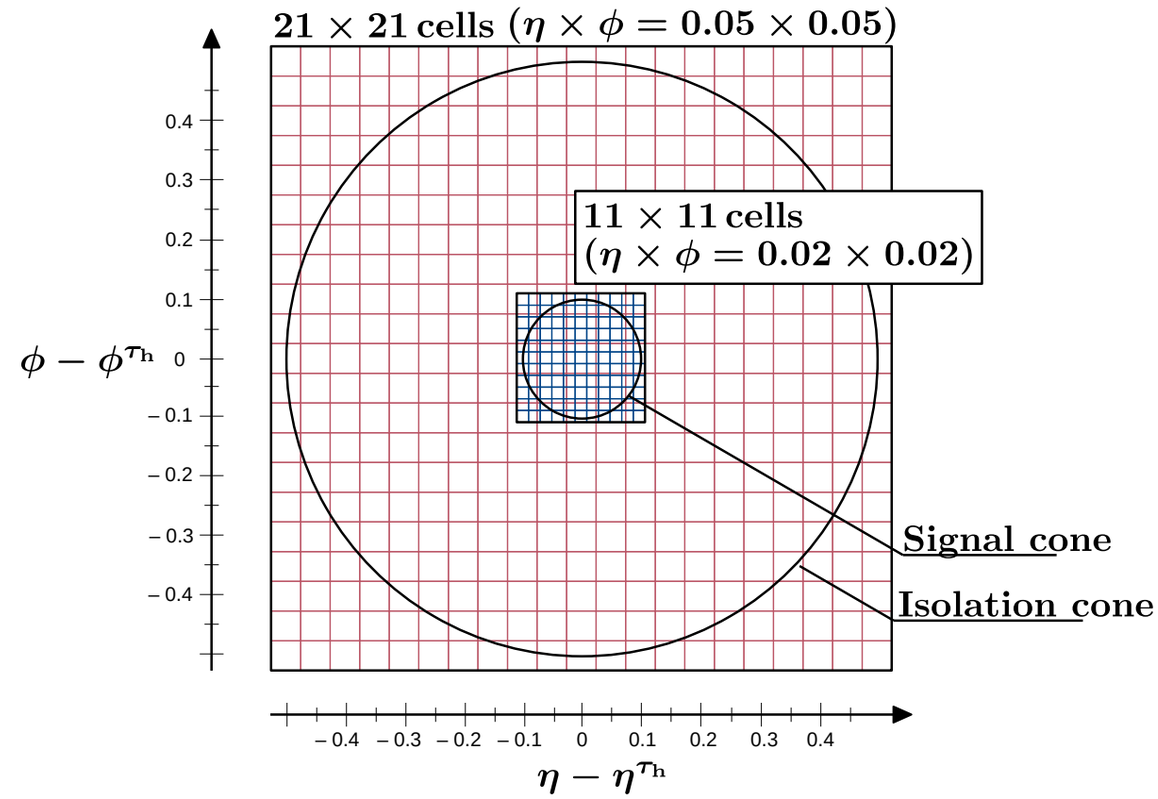


QCD jet (background)

(*) See JINST 13 (2018) P10005 for more details

DNN-based tau ID

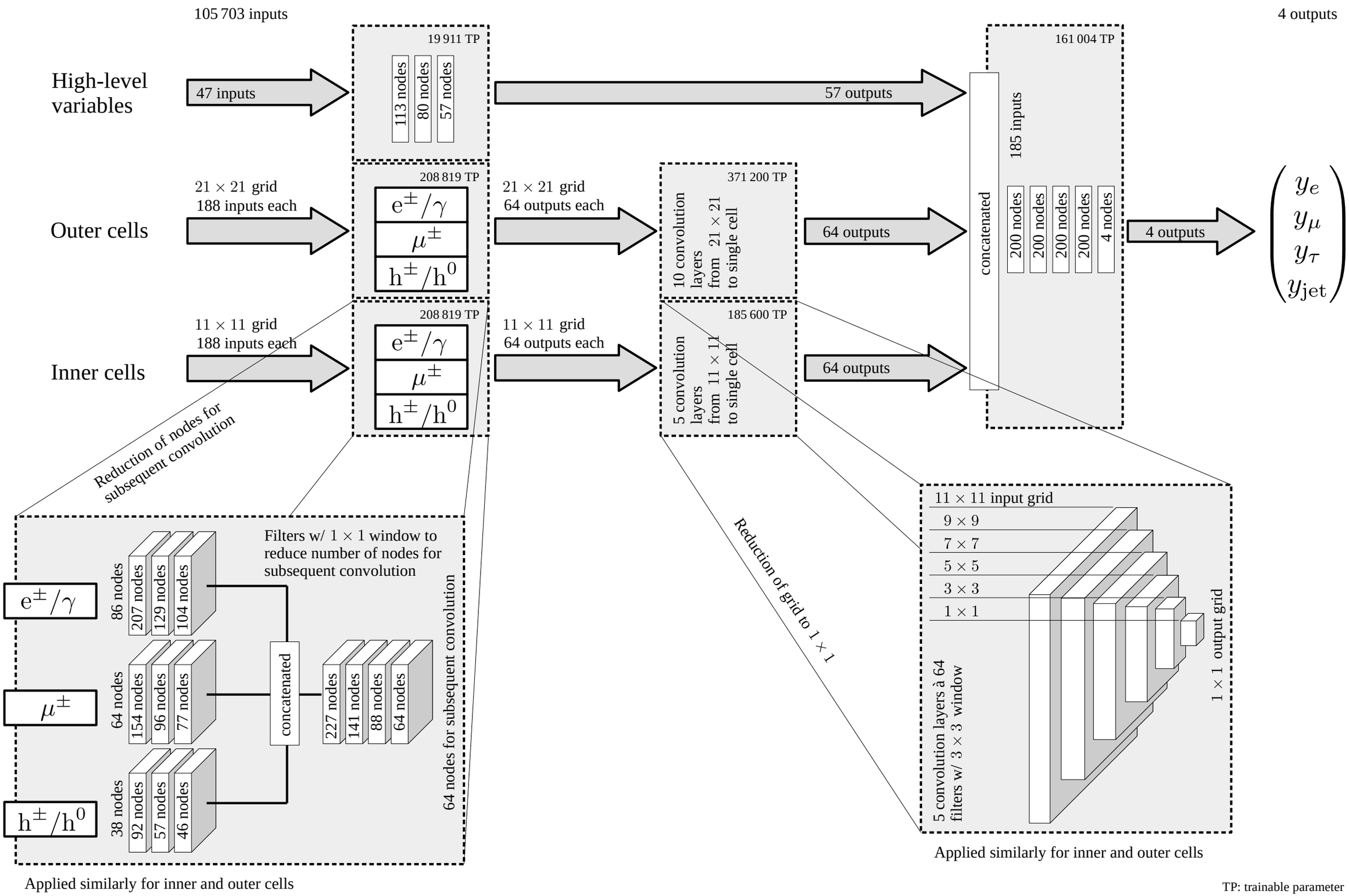
- To efficiently discriminate τ_h against main backgrounds, the detailed information from multiple sub-detectors within CMS must be exploited, including the **inner tracker**, the electromagnetic (**ECAL**) and hadronic (**HCAL**) calorimeters, and the **muon chambers**
- **DeepTau** is a multiclass tau identification algorithm based on a convolutional deep neural network (DNN) [1]
- The training is performed on a balanced mix of ≈ 140 million τ_e, τ_μ, τ_h **and** τ_j **candidates** coming from Drell-Yan, $t\bar{t}$, W+jets and Z' 2017 MC samples
- **Particle** = PF candidate *OR* fully reconstructed electron *OR* fully reconstructed muon
- Particles belonging to the signal and isolation cones are split into two $\eta \times \varphi$ grids



Particle	N_{var}
PF charged hadron	27
PF neutral hadron	7
Electron	37
PF electron	22
PF photon	23
Muon	37
PF muon	23

[1] Submitted to JINST. Preprint at arXiv:2201.08458 (2022)

DeepTau architecture



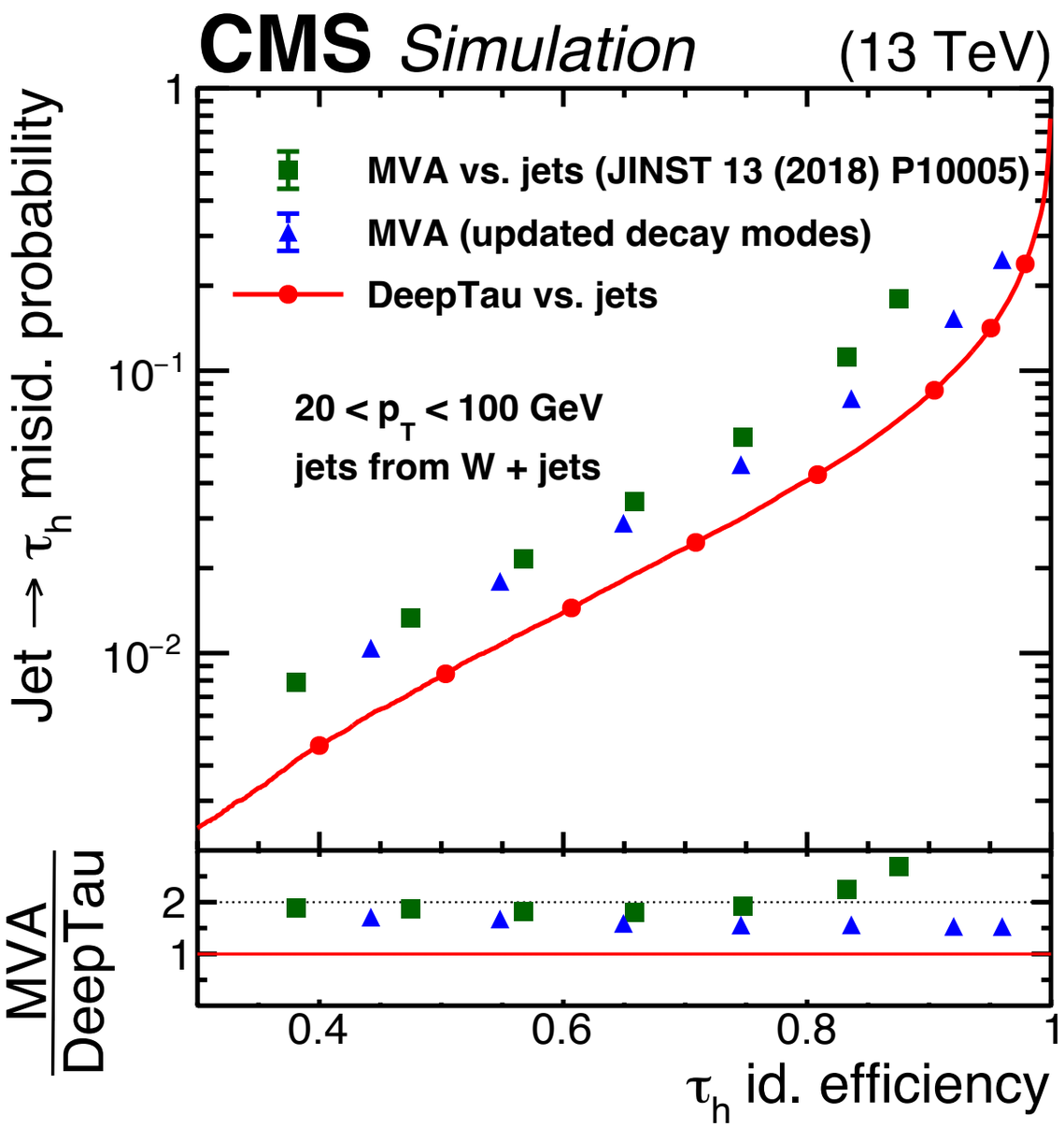
DeepTau: final discriminator and minimisation

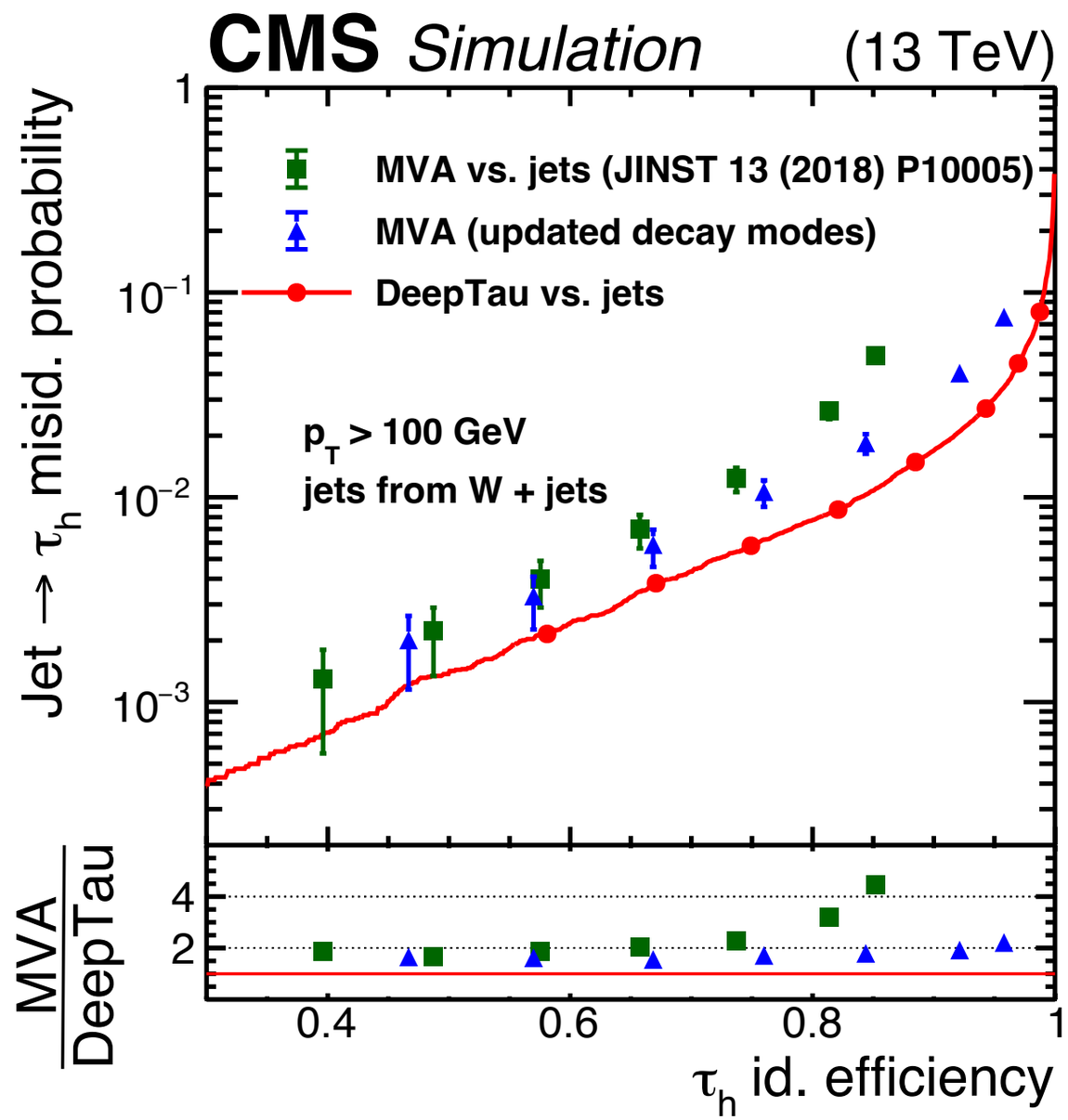
- The final discriminator is chosen to be

$$D_\tau^\alpha(\boldsymbol{p}) = \frac{p_\tau}{p_\tau + p_\alpha}, \quad where\ \alpha \in \{e, \mu, j\}$$

- A **custom loss function** based on the focal loss is defined in order to ensure the best performance for a wide tau ID efficiency range
- The loss function is minimised using **NAdam** algorithm (Adam with Nesterov momentum)
- The training is run for **10 epochs** on GeForce RTX 2080
  - Average speed $\approx 3\ days/epoch$
- The best performance on the validation set is achieved after **7 epochs**
  - The corresponding NN is chosen as the final

DeepTau discrimination against jets



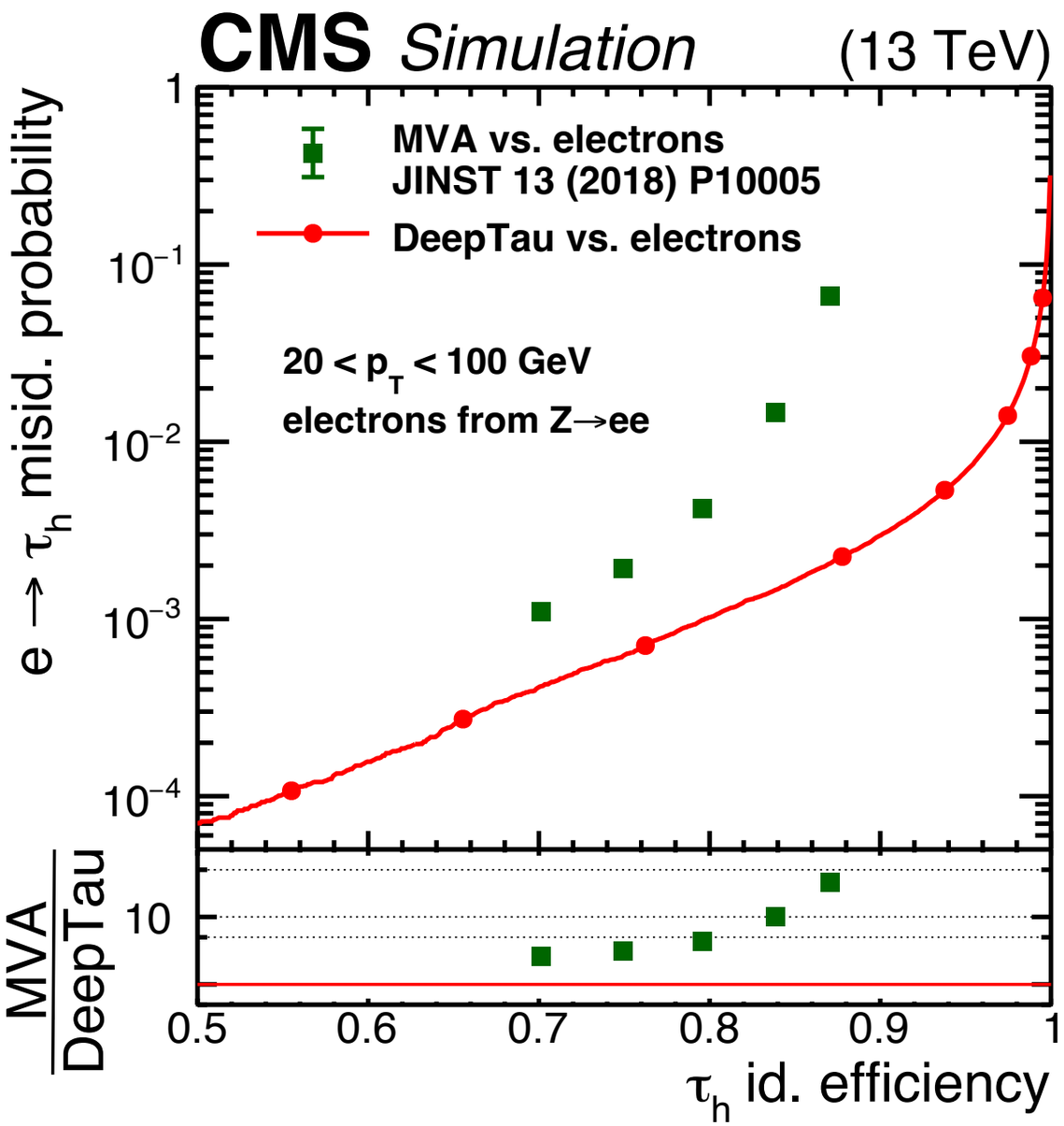


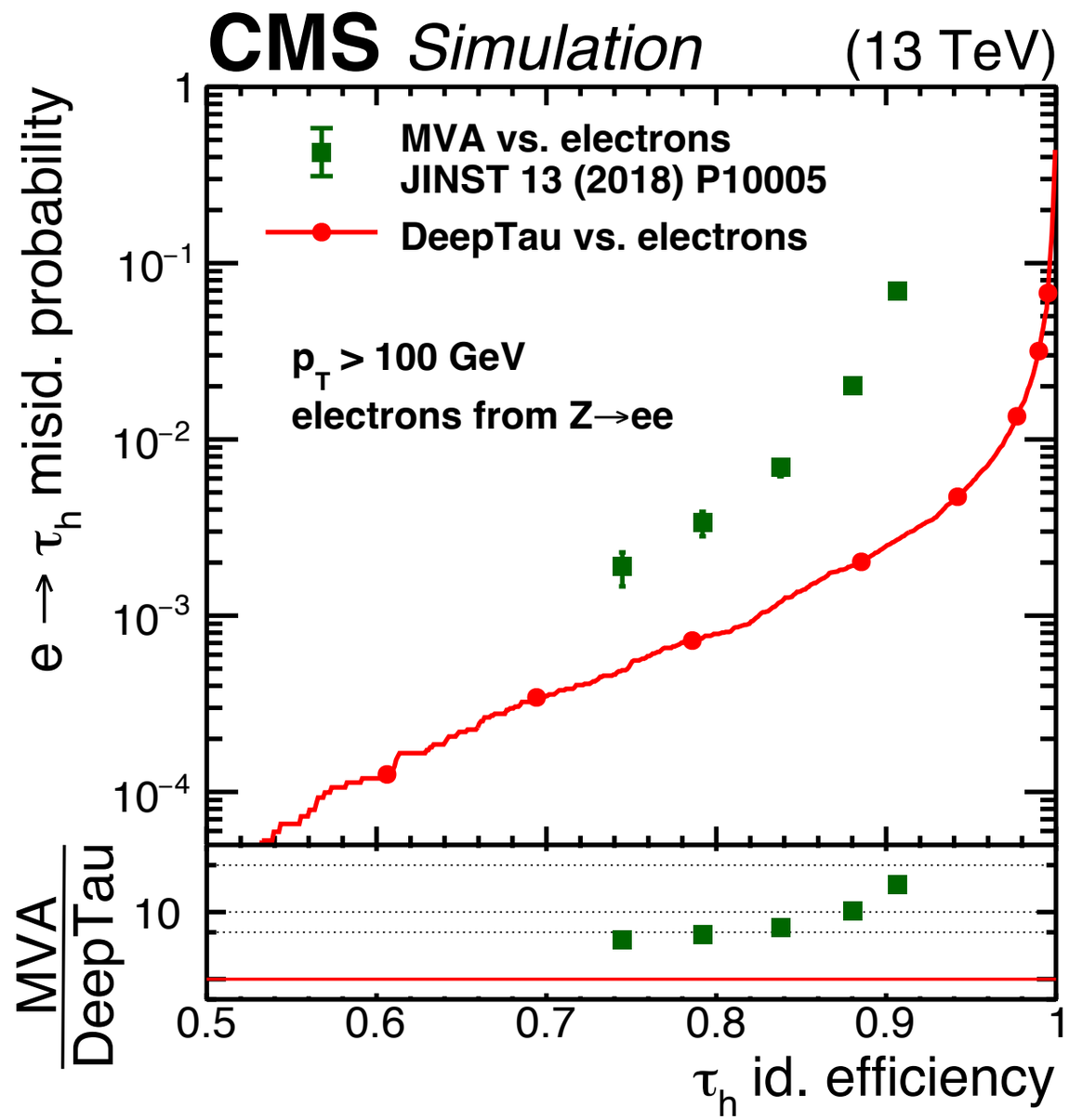
- Consistent improvement at both **low and high** p_T
- Jet misidentification probability **reduces by ~50..80%**

τ_h from $H \to \tau\tau$ MC
τ_j from W+jets MC

Working points of the discriminators are indicated by markers

DeepTau discrimination against electrons



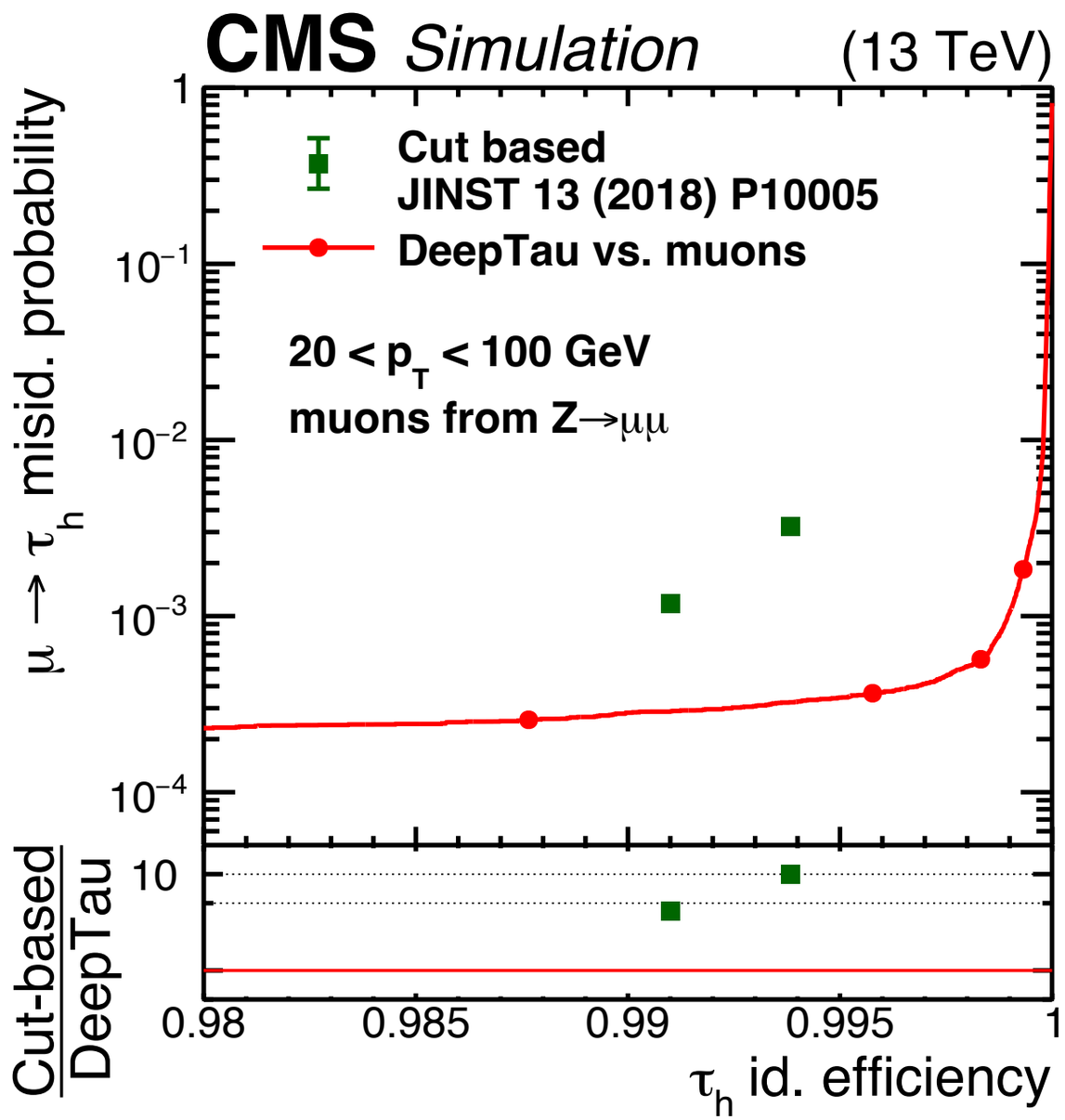


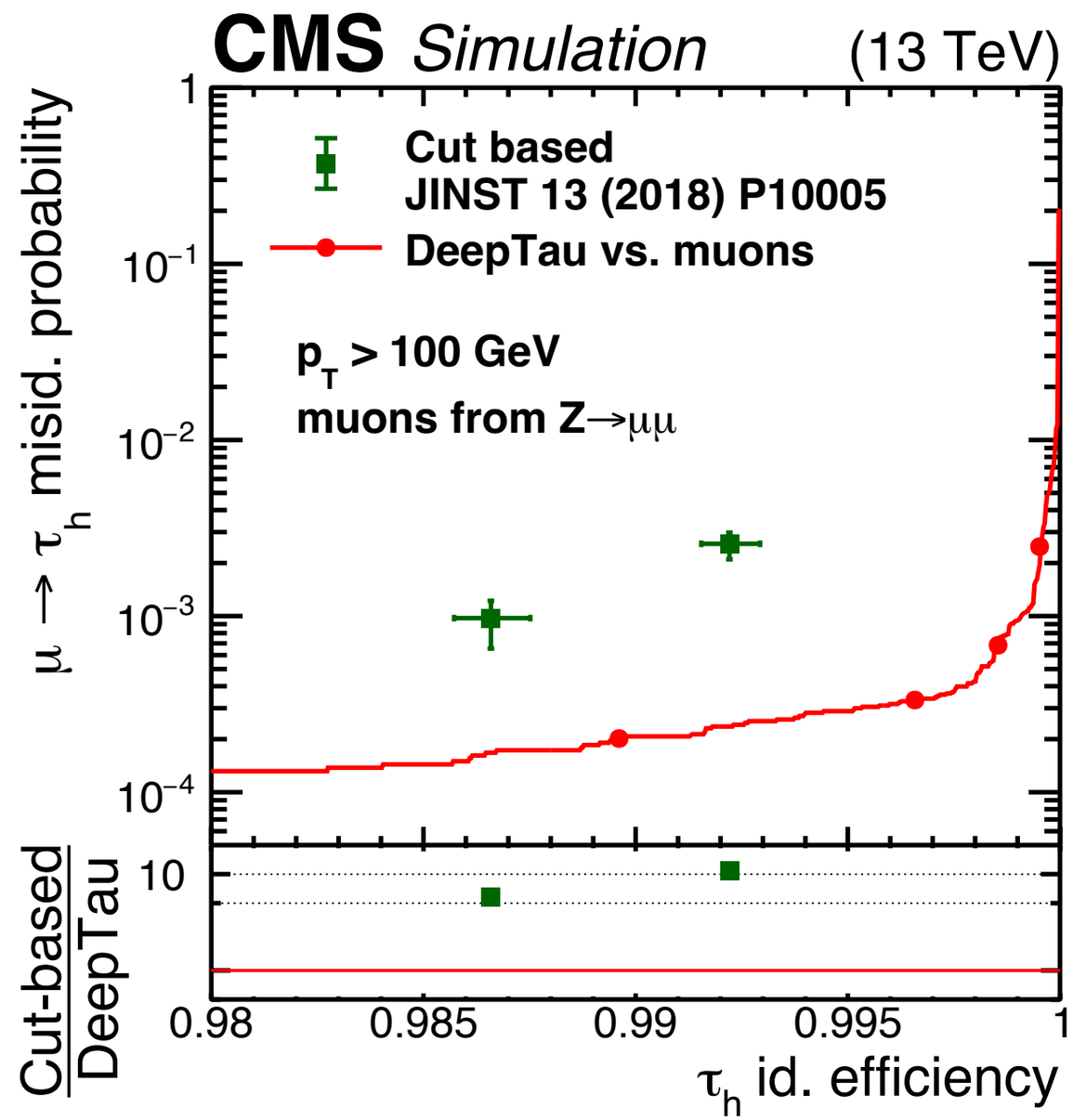
- Consistent improvement at both **low and high** p_T
- Electron misidentification probability **reduces by ~60..95%**

τ_h from $H \to \tau\tau$ MC
τ_e from $Z \to ee$ MC

Working points of the discriminators are indicated by markers

DeepTau discrimination against muons



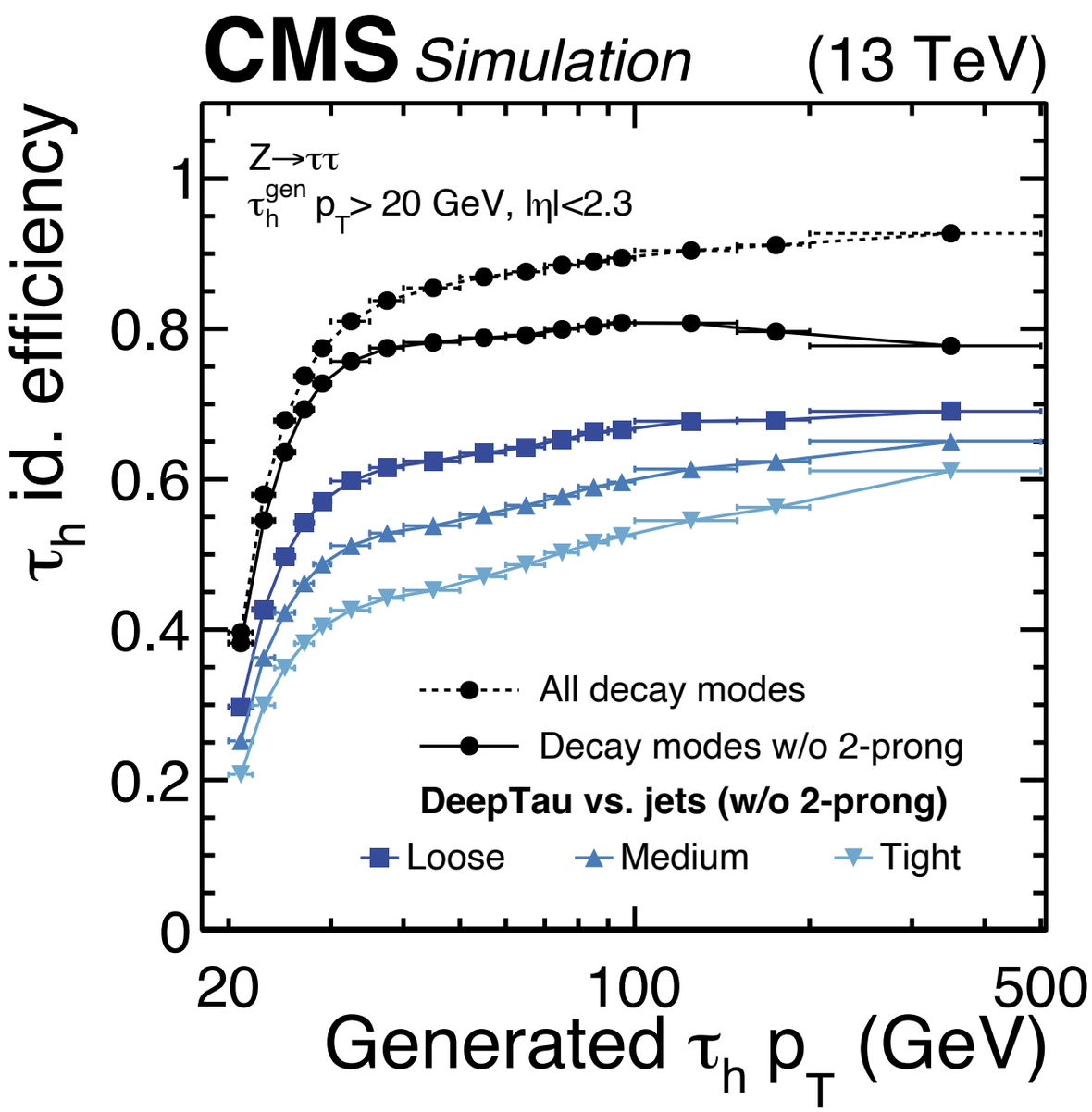


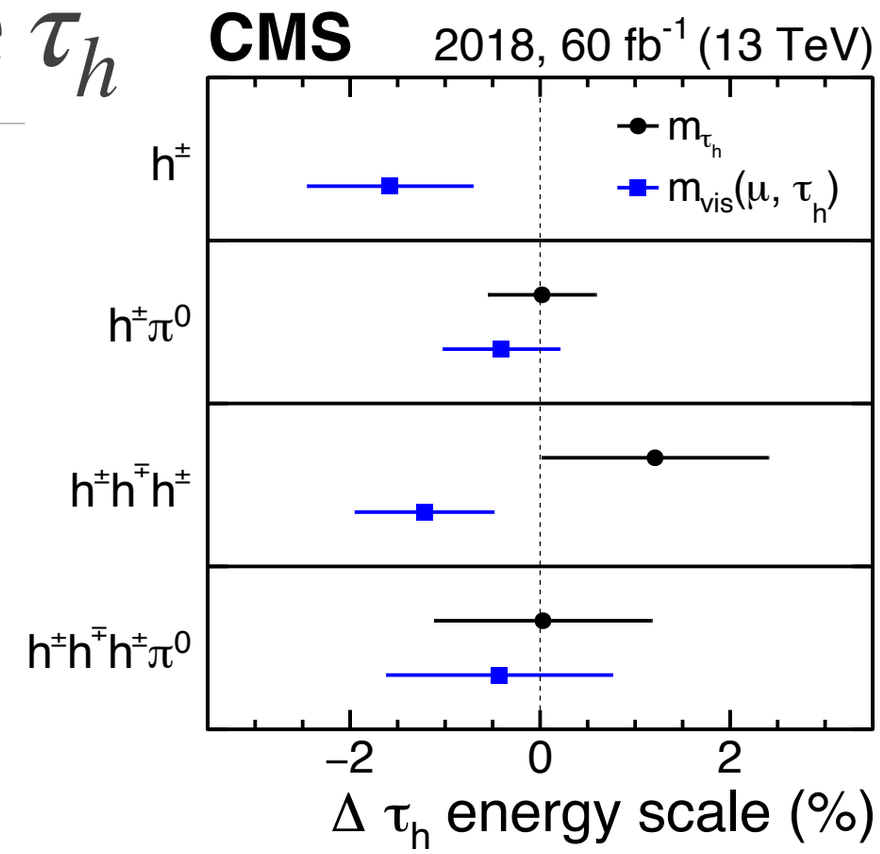
- Consistent improvement at both **low and high** p_T
- Muon misidentification probability **reduces by ~65..90%**

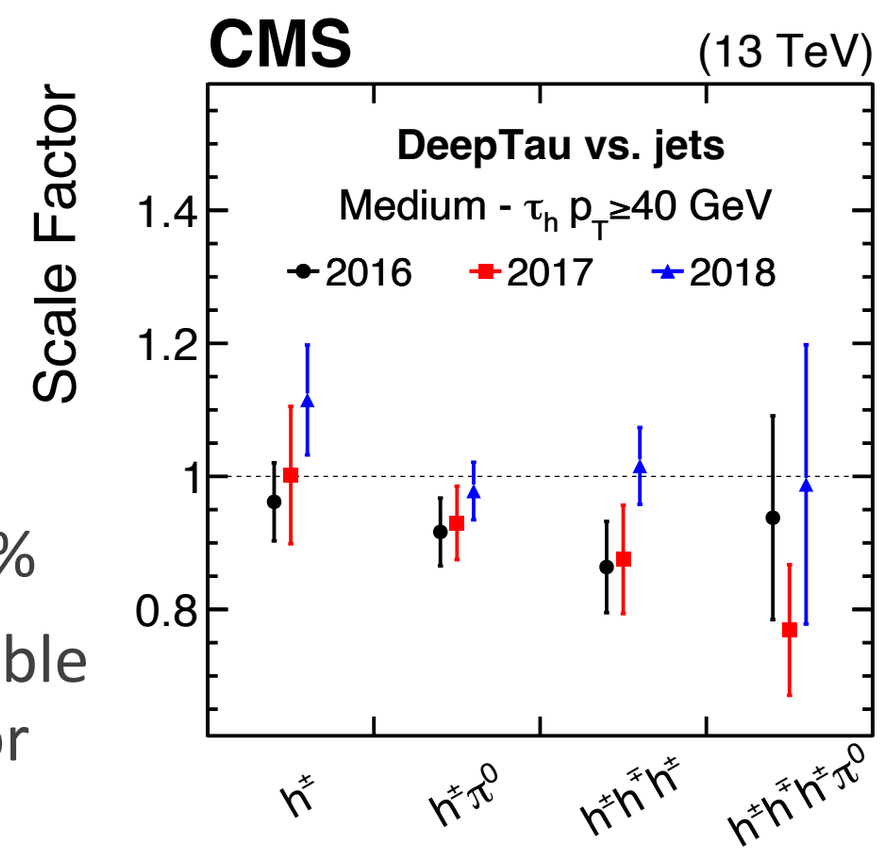
τ_h from $H \to \tau\tau$ MC
τ_μ from $Z \to \mu\mu$ MC

Working points of the discriminators are indicated by markers

DeepTau performance for genuine τ_h



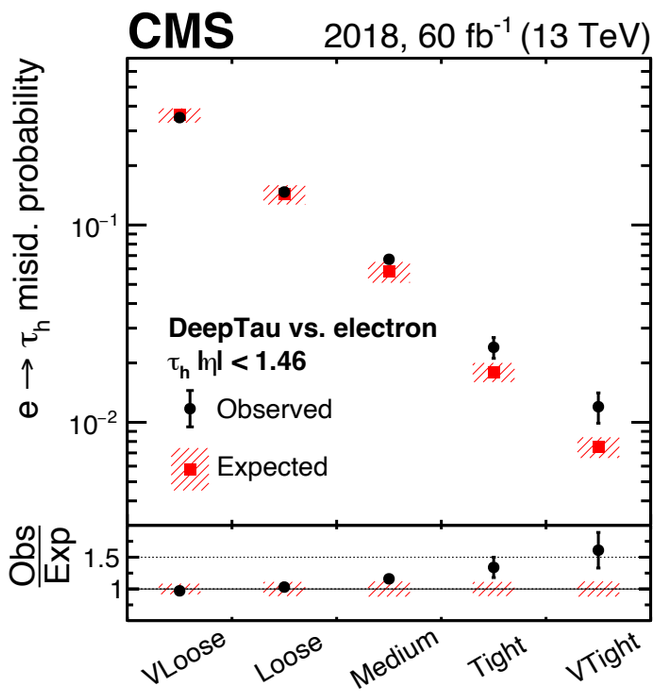


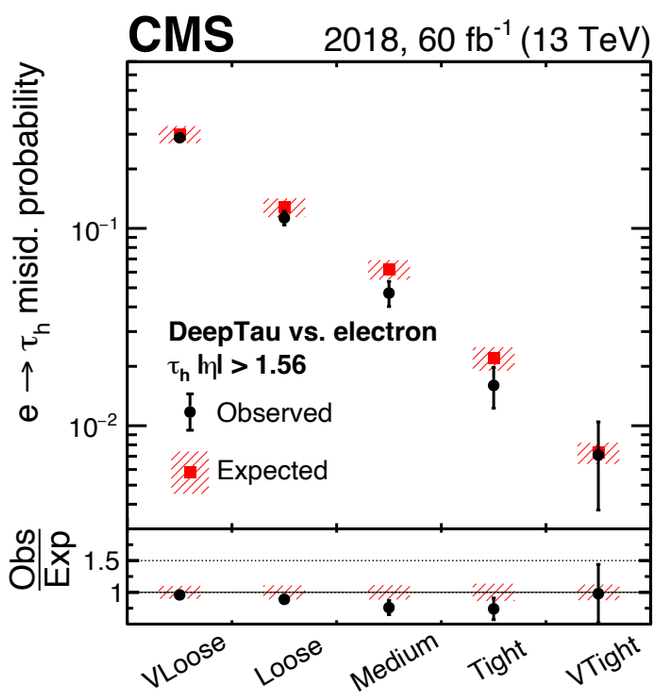


- Tau efficiency is about 49% at $jet \rightarrow \tau_h$ misid rate ~0.5%
- Tau energy scale and efficiencies are generally compatible within a few percents between data and simulations for most commonly used working points

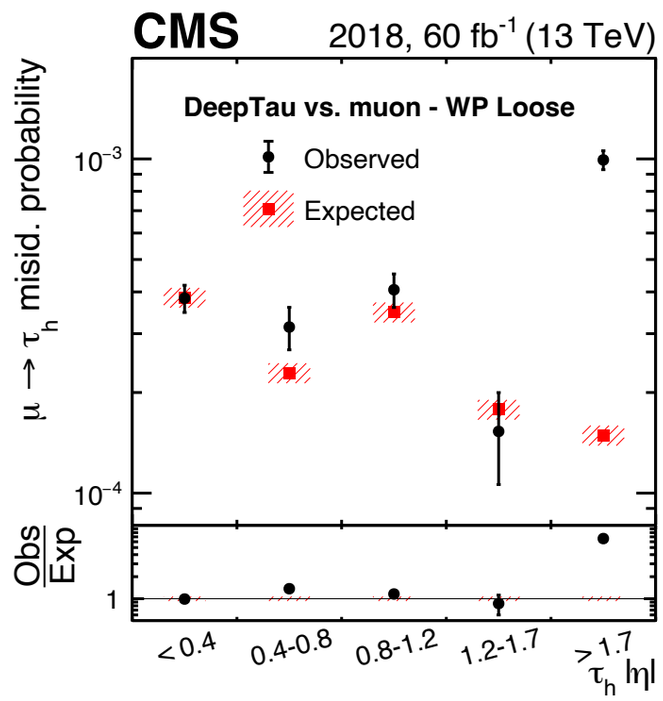
DeepTau performance for fake taus

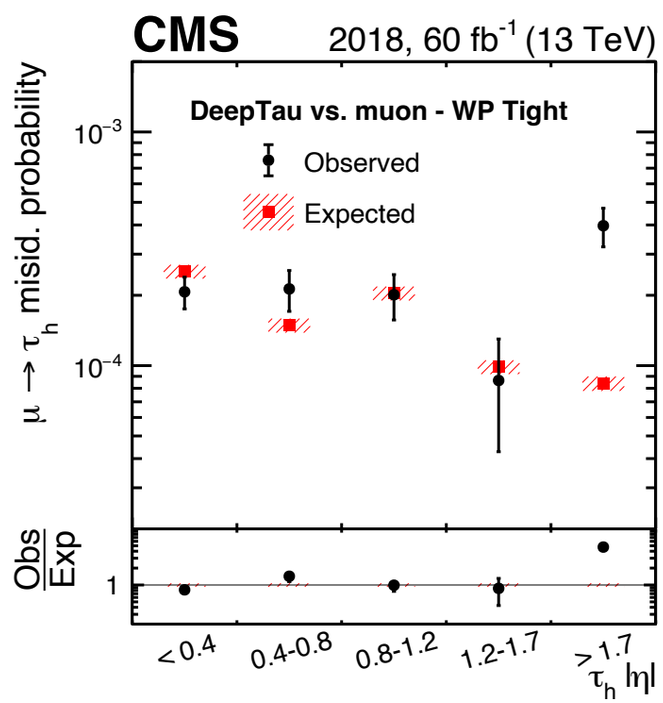
$e \rightarrow \tau_h$



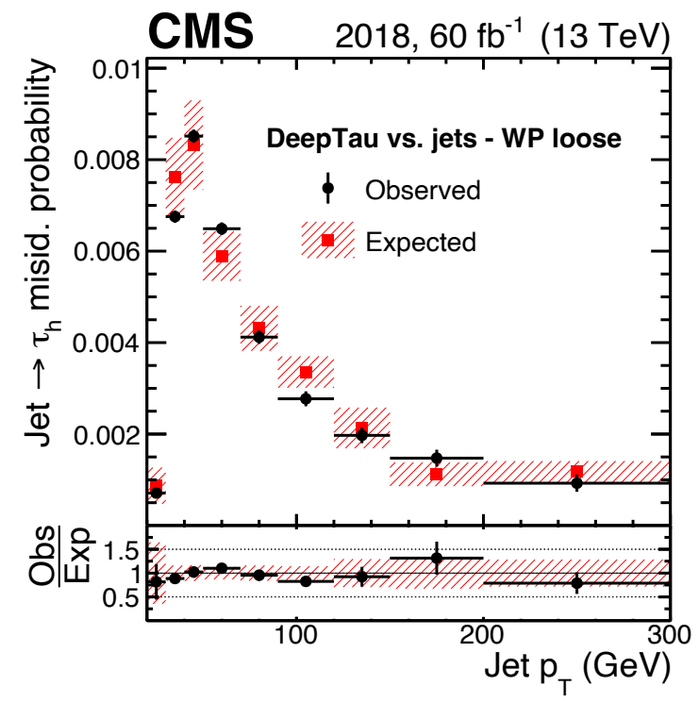


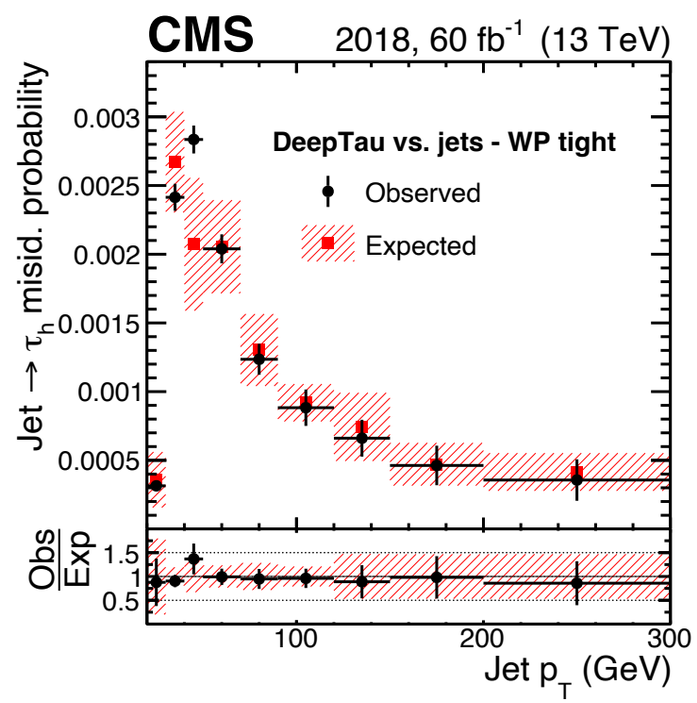
$\mu \rightarrow \tau_h$





$jet \rightarrow \tau_h$





Good modelling of fake τs in the most parts of the parameter space

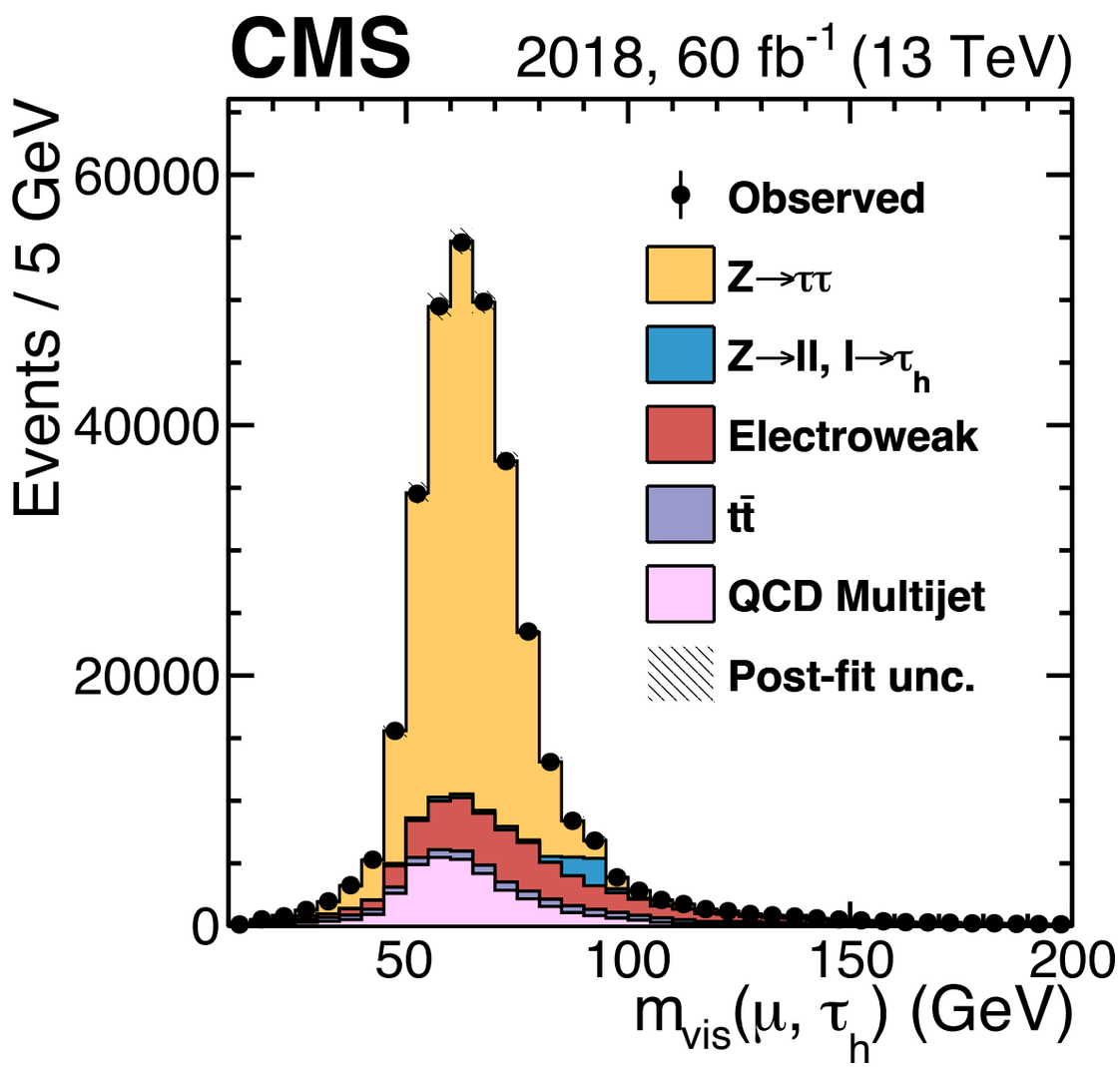
Data-MC control plots

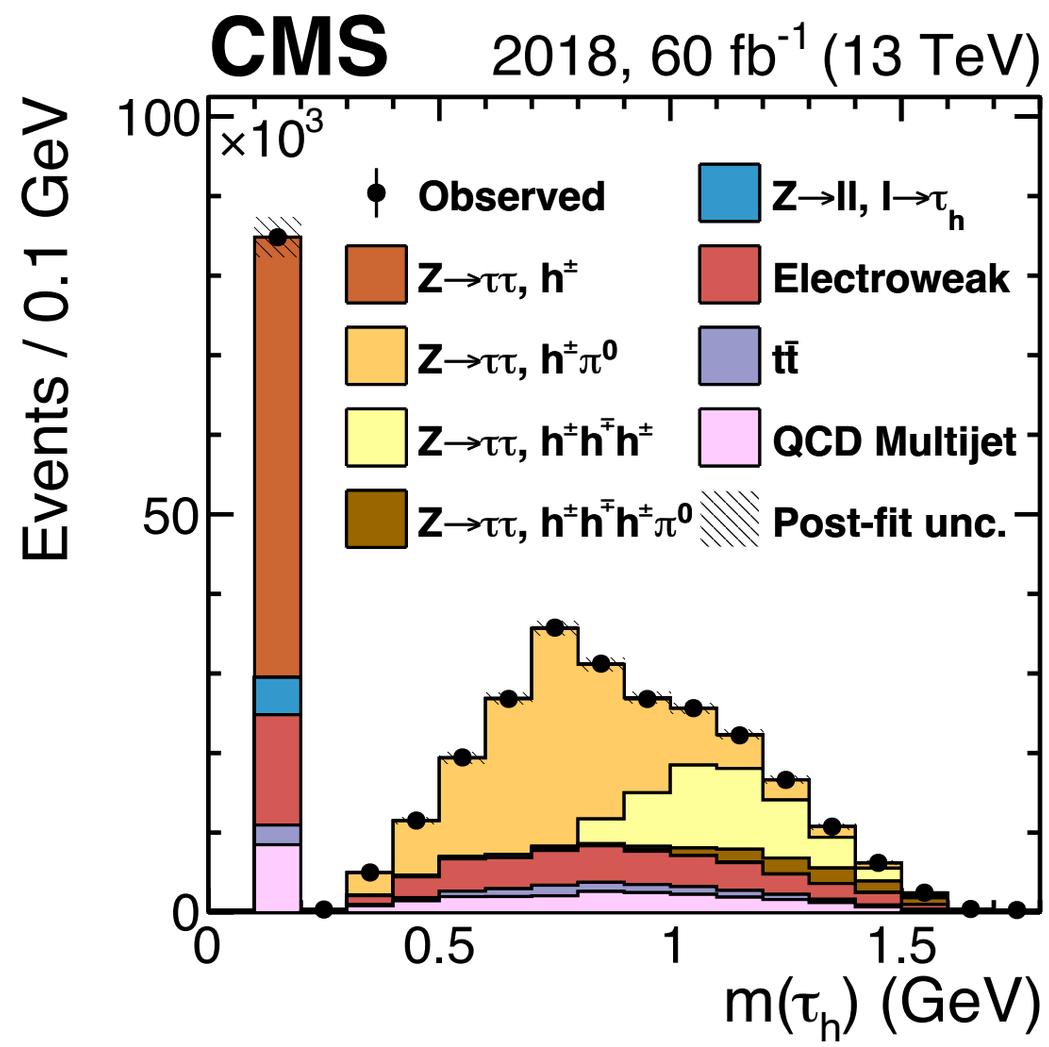
Using minimal preselection on muon and tau candidates:

- Well identified and isolated muon with $p_T > 25$ GeV, $|\eta| < 2.4$, $|dz| < 0.2$ cm
- Tau candidates with $p_T > 20$ GeV, $|\eta| < 2.3$, $|dz| < 0.2$ cm
- $\mu\tau$ pair with an opposite charge and $\Delta R(\mu, \tau) > 0.5$, $M_T < 60$ GeV, $|\Delta\eta| < 1.5$

Selection using DeepTau IDs:

- Tight WP against jets
- VVLoose WP against electrons
- VLoose against muons





In both plots modelled contributions are fit to the data

Conclusions

- The DNN-based algorithm, **DeepTau**, to discriminate hadronic tau decays from the main background sources has been presented [1]
- Introduction of DeepTau ID provides considerable improvement in the performance in tau identification for Run 2
  - Jet mis-id probability **reduces by more than 50%**
  - Reduction of mis-id probability for electrons (muons) is **up to 95% (90%)**
- DeepTau has already been used in several recent CMS physics analyses with Run 2 data
- Further improvement of the algorithm are ongoing for Run 3:
  - DeepTau for online event selection
  - Use of domain adaptation techniques to improve the modelling
  - Other NN architectures (e.g. graph neural networks)

[1] Submitted to JINST. Preprint at arXiv:2201.08458 (2022)