ICHEP 2020 | PRAGUE

40th INTERNATIONAL CONFERENCE ON HIGH ENERGY PHYSICS

VIRTUAL CONFERENCE

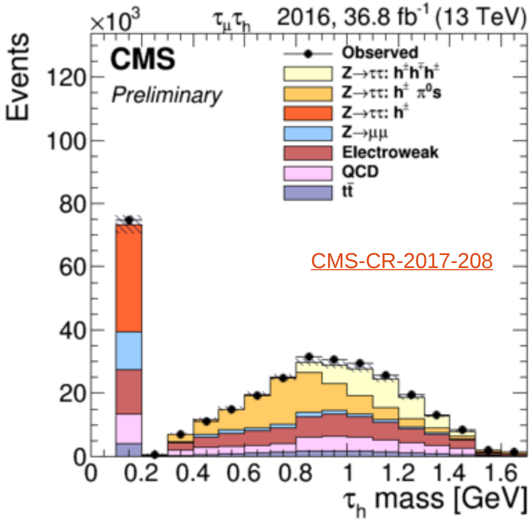
Tau identification exploiting deep learning techniques

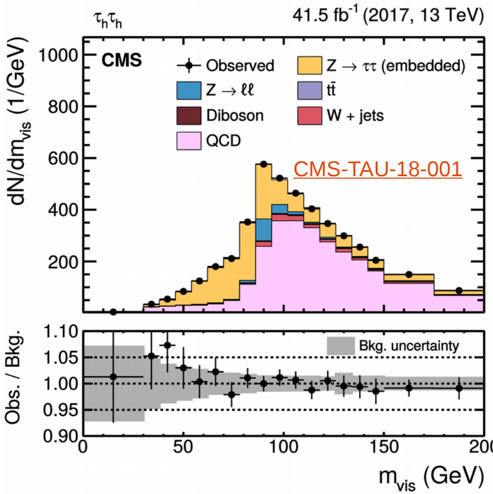
Andrea Cardini

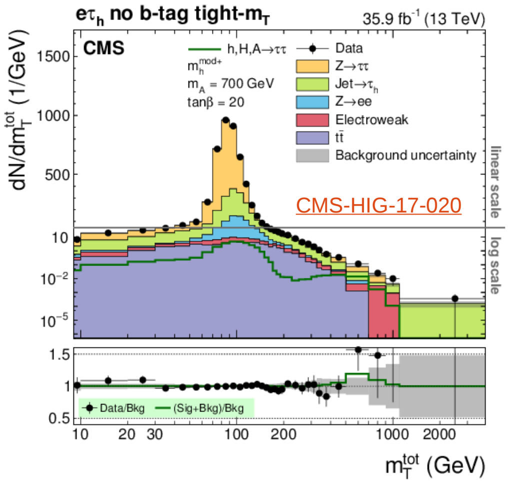
on behalf of the CMS Collaboration

Tau physics

- τ leptons (taus) are the heaviest leptons in the SM
- They can be used for several measurements with final states involving taus:
  - Test EWK interaction
  - Yukawa couplings of H → fermions
  - Study of the CP properties of the Higgs
  - Tau polarization in Z boson decays
- Searches for BSM physics:
  - Leptoquarks, SUSY, high mass resonances

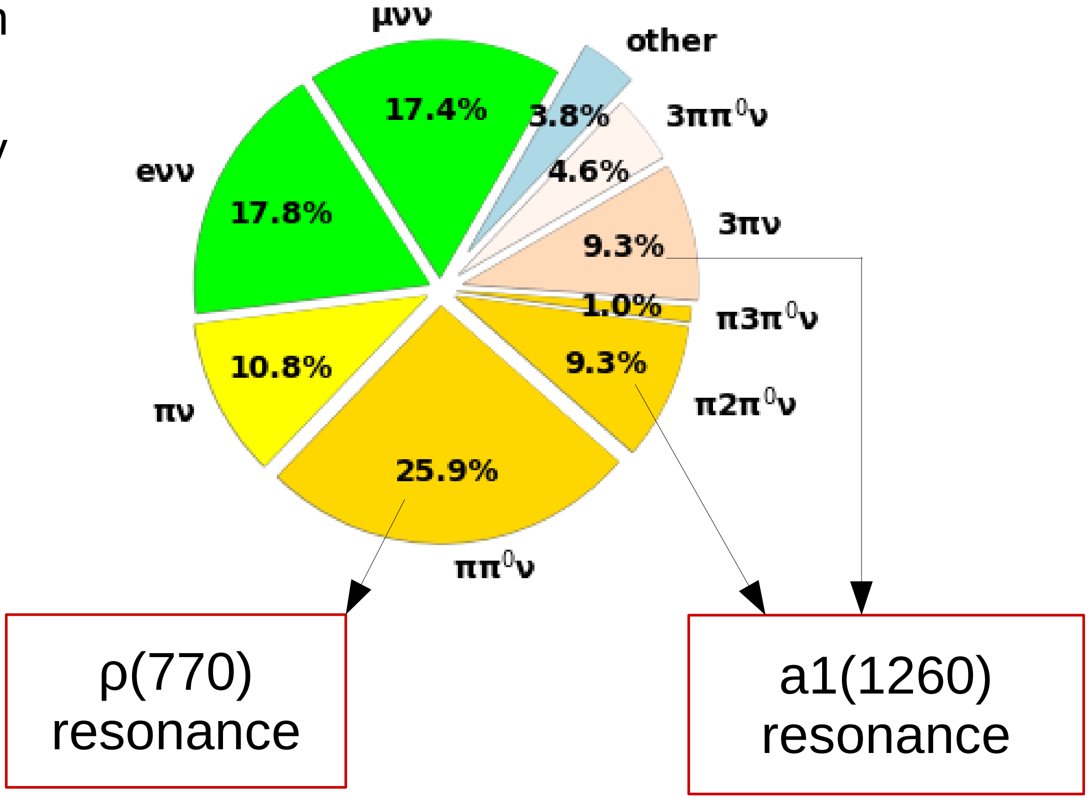




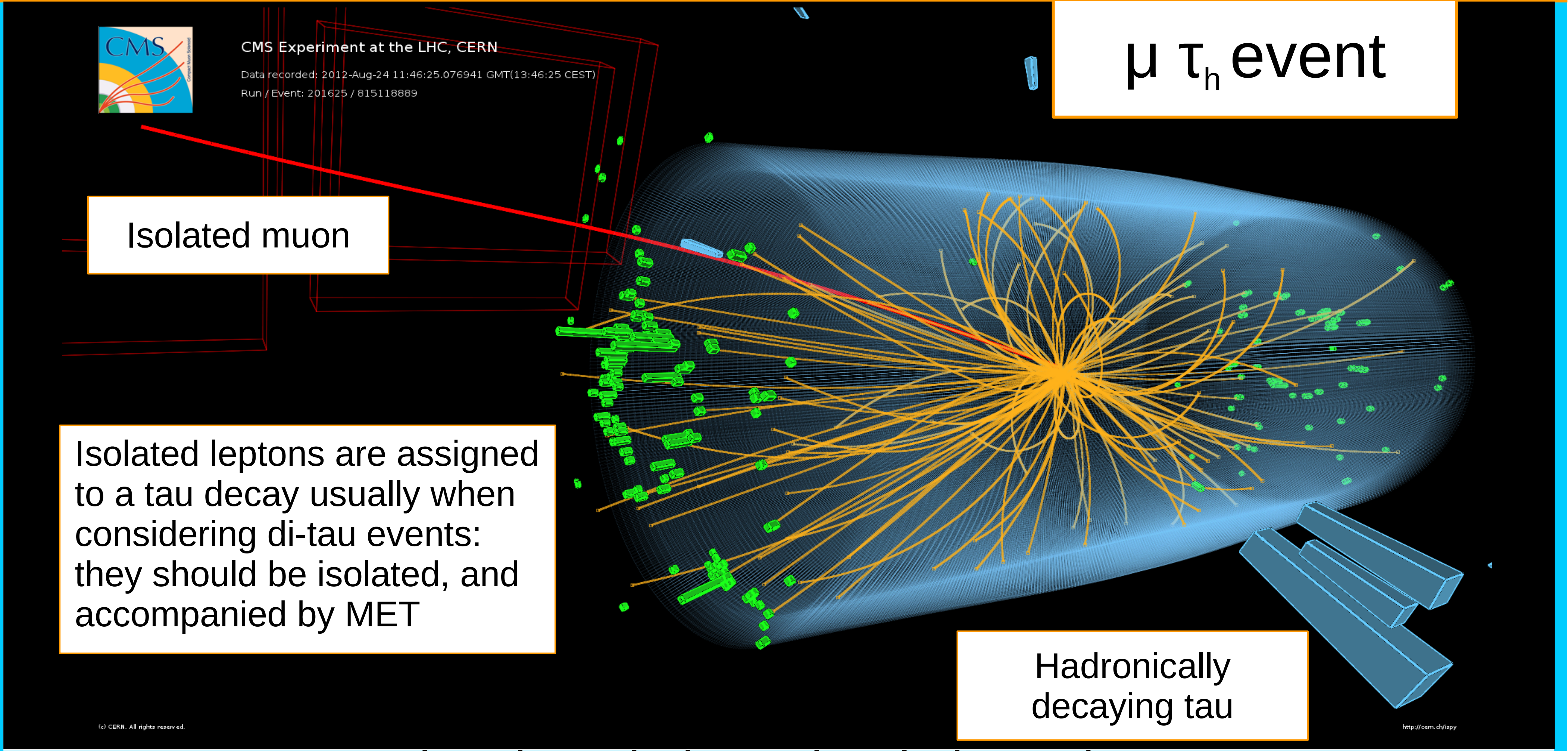


Tau properties and decays

- Tau properties:
  - mass ~1.78 GeV → is the only lepton that can decay hadronically
  - average lifetime ~3 x 10^{-13} s → decay length of ~1.5 mm (with E ~30 GeV)
- Tau decays involve charged particles → **prongs**
- **Leptonic decays** always involve only one prong
- **Hadronic decays** are mostly via intermediate **mesonic resonances** and can involve 1 or 3 prongs



Di tau event



$\mu\ \tau_h$ event

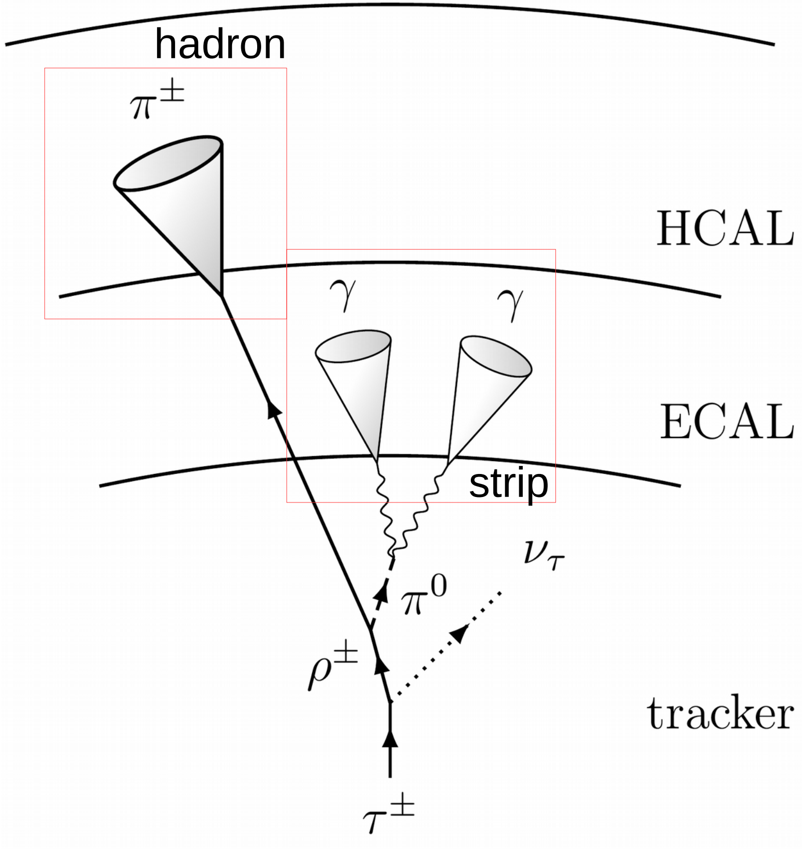
Isolated muon

Isolated leptons are assigned to a tau decay usually when considering di-tau events: they should be isolated, and accompanied by MET

Hadronically decaying tau

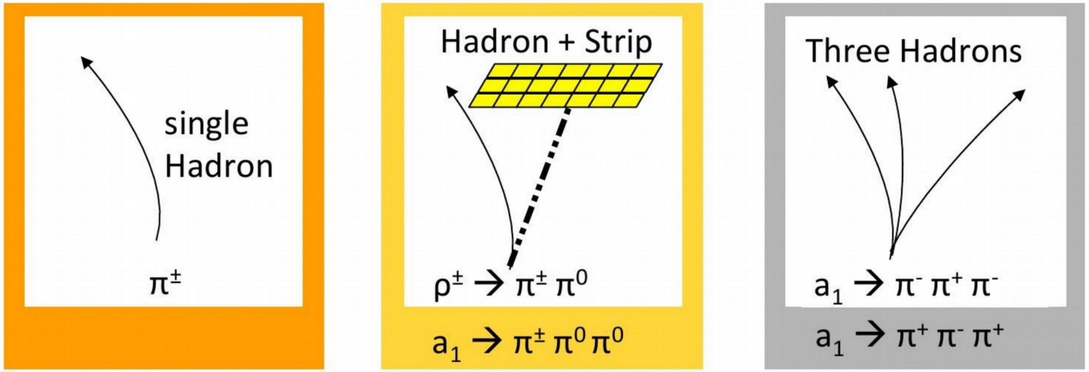
PF and HPS algorithms

credits to Izaak W. Neutelings for the image



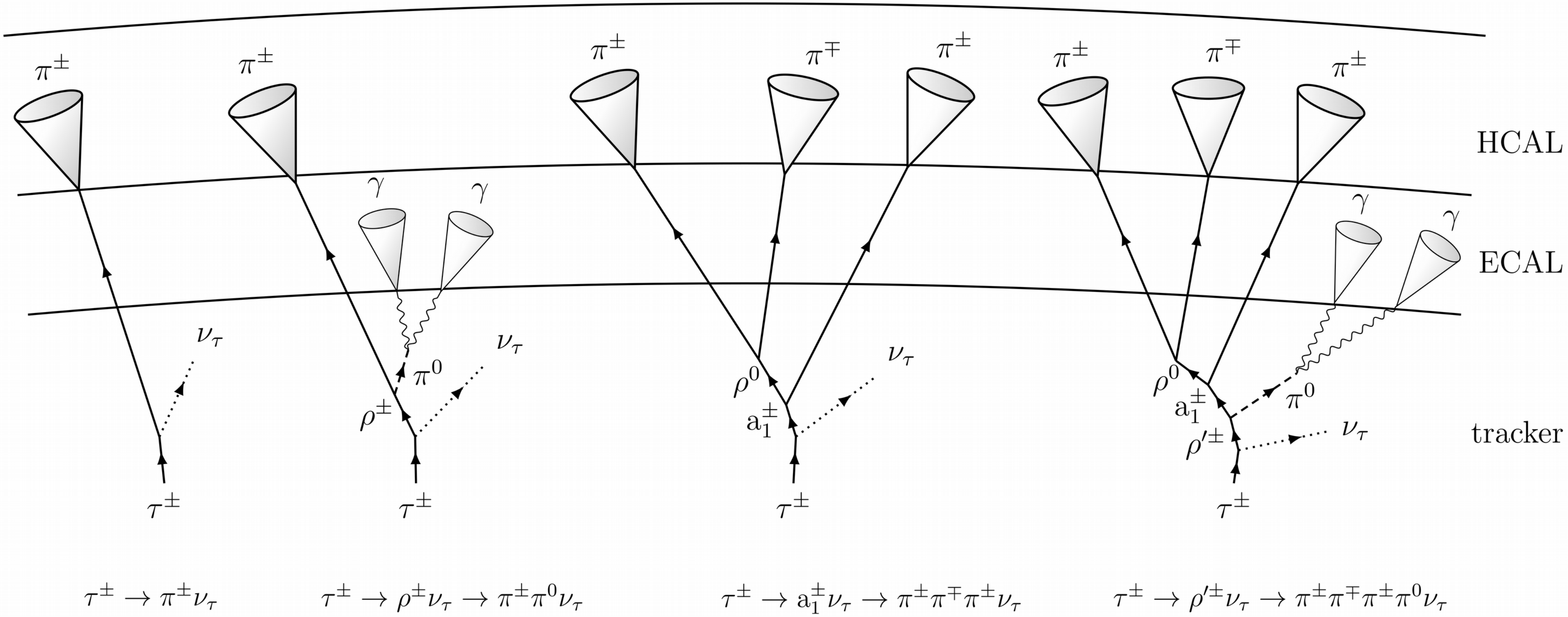
$$\tau^\pm \to \rho^\pm \nu_\tau \to \pi^\pm \pi^0 \nu_\tau$$

- All objects in CMS are reconstructed via the Particle Flow (PF) algorithm → PFCandidates
- The Hadron-plus-strip (HPS) algorithm combines:
  - PFCandidates for jets → hadronic jets
  - EM showers in ECAL which are elongated in φ → strips
- 3 reconstructed decay channels in older version → now 4 decay modes are identified



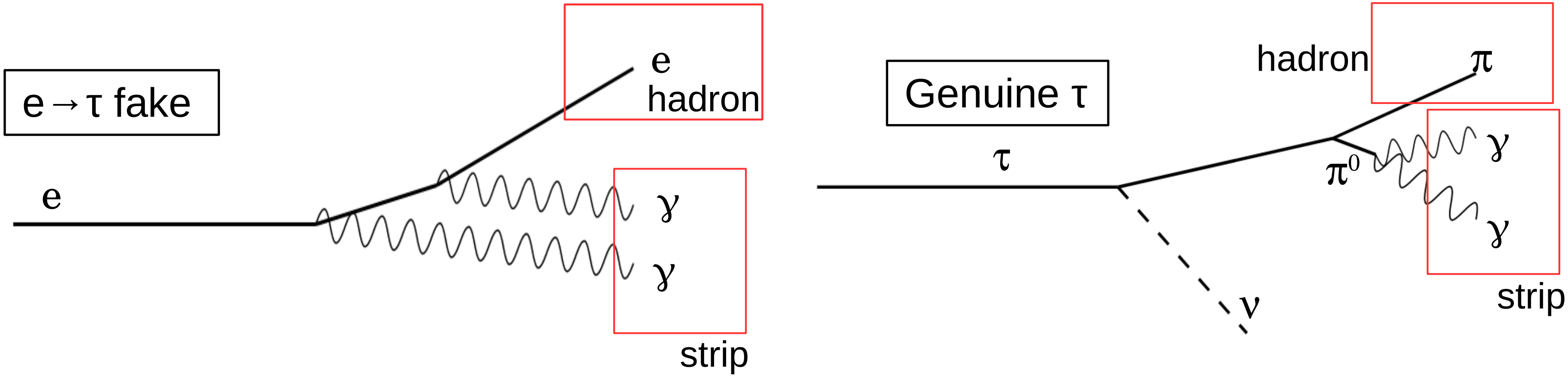
Hadronic tau reconstructed decays

credits to Izaak W. Neutelings for the image



Misidentification as hadronic taus

- Several objects can be misidentified as hadronic taus by the HPS algorithm:
  - **Jets** → a highly collimated quark or gluon jet can be mistaken for any tau decay
  - **Muons** → mainly affects the 1 prong channel
  - **Electrons** → can emit photons via bremsstrahlung radiation and mimic the ρ decay



- The misidentification is reduced via the DeepTau neural network based algorithm

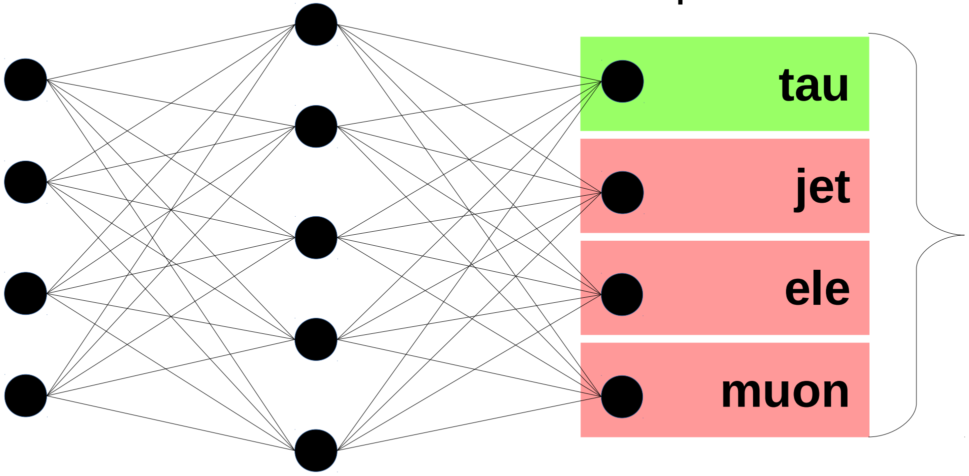
The DeepTau Convolutional NN

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

Inputs:

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

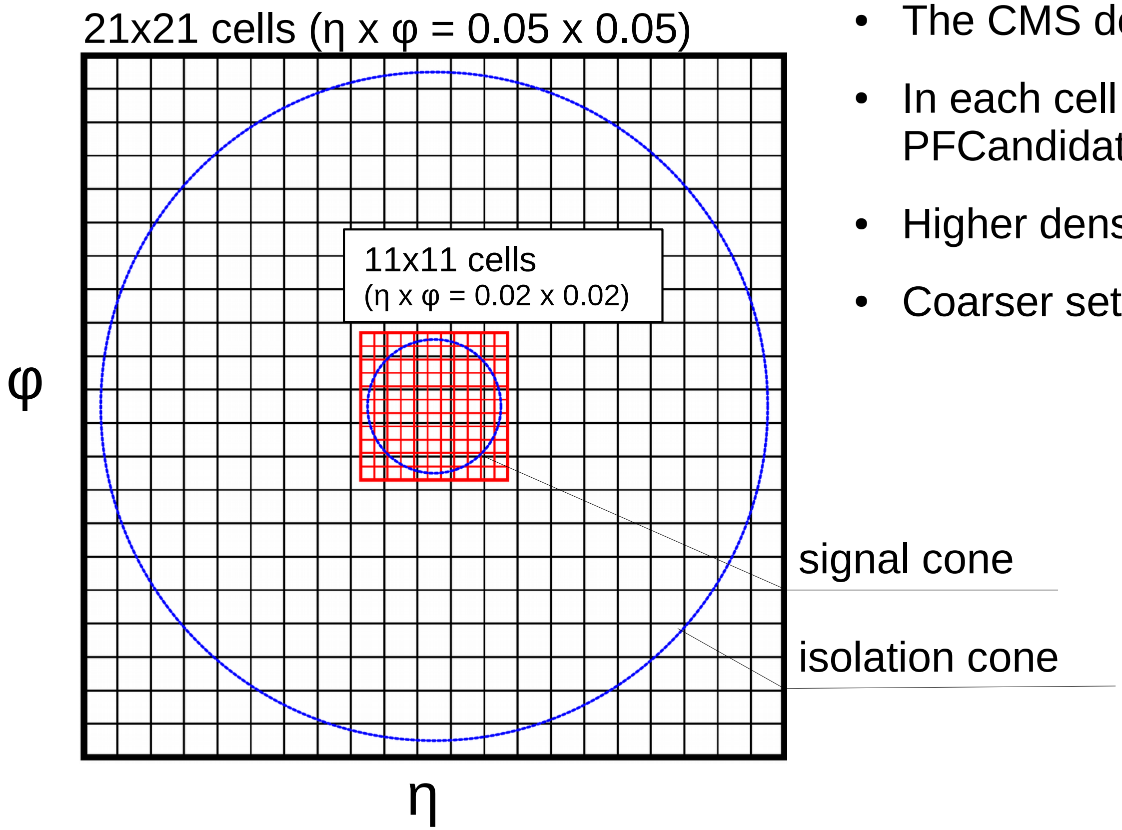
4 output nodes



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

Classifiers used to reduce the miss-ID rate

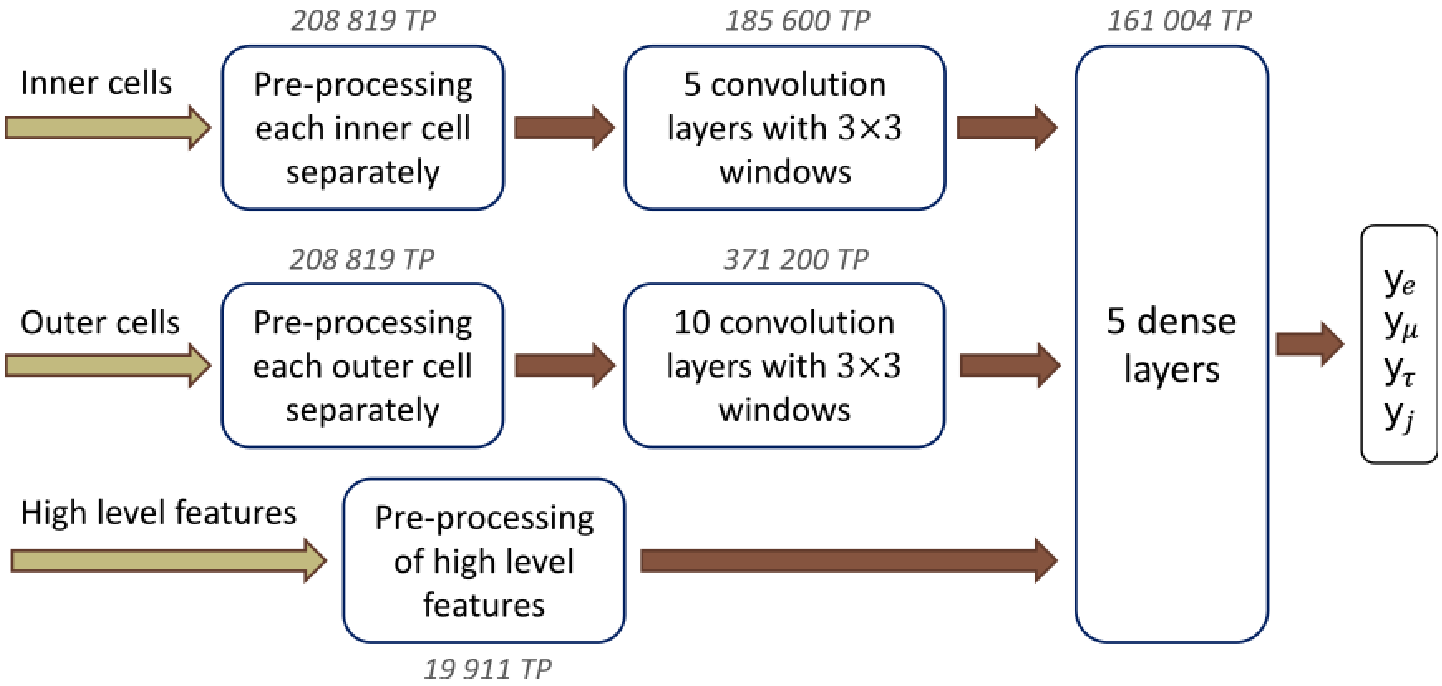
Input features



- Low level inputs are based on the tau decay products
- The CMS detector is divided in cells of η x φ
- In each cell all available information for the leading PFCandidate is stored
- 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, HPS-DM, charge, IP, number of charged prongs and neutral constituents, etc.
        - Average energy in the event, Δη ECAL, etc.

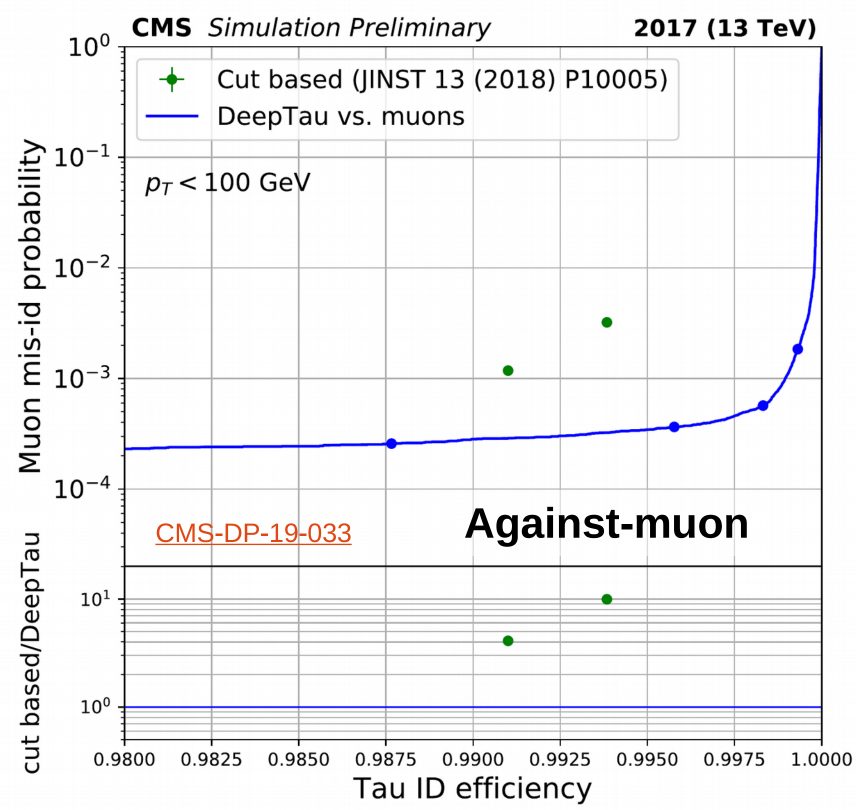
The NN structure and training

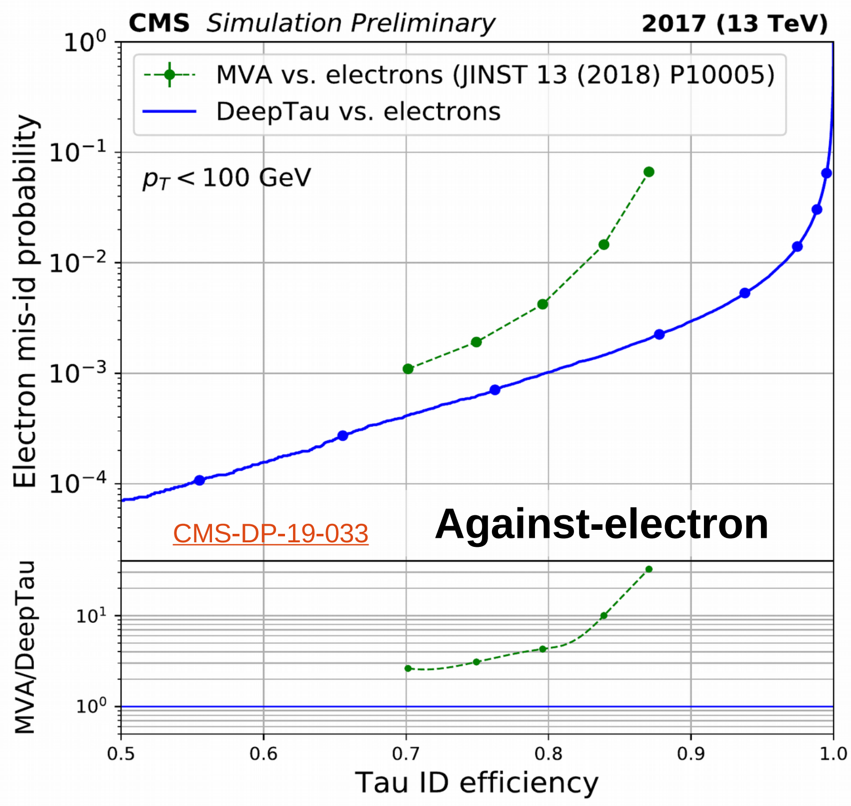
- The convolutional NN takes as inputs $\mathcal{O}$(100k) low and high-level features and has $\mathcal{O}$(1.5M) trainable parameters (TP)
- Low-level features are pre-processed using 3 convolutional layers to reduce the number of features



- The training is performed using the NAdam algorithm
- Due to the number of parameters the training is performed on one GPU and takes ~3 days/epoch

The against-lepton classifiers

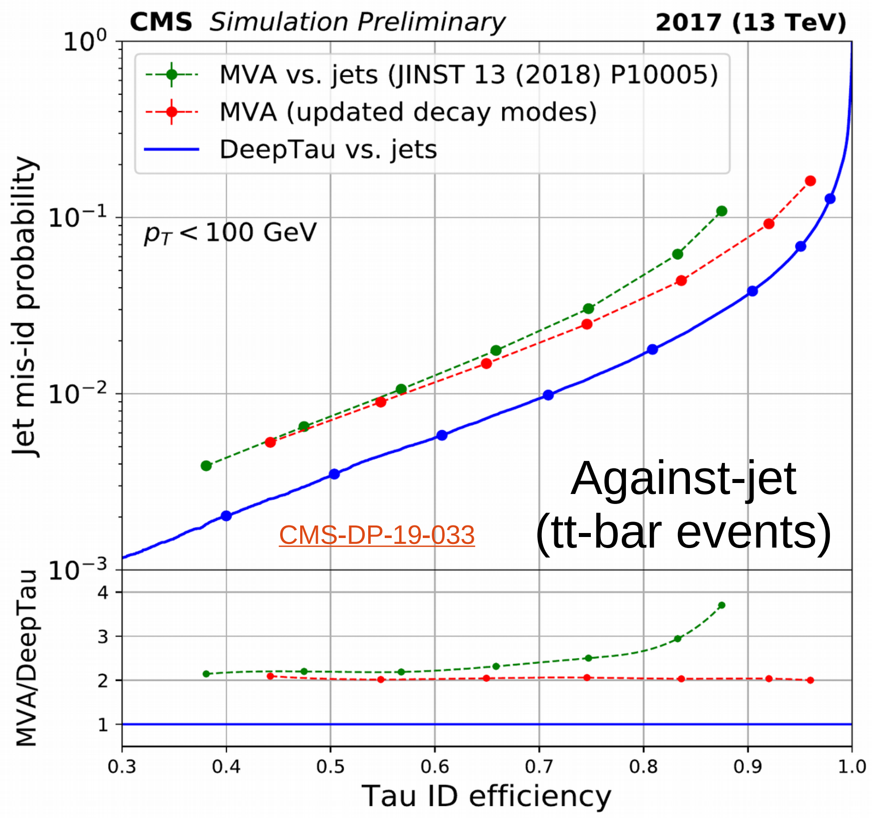


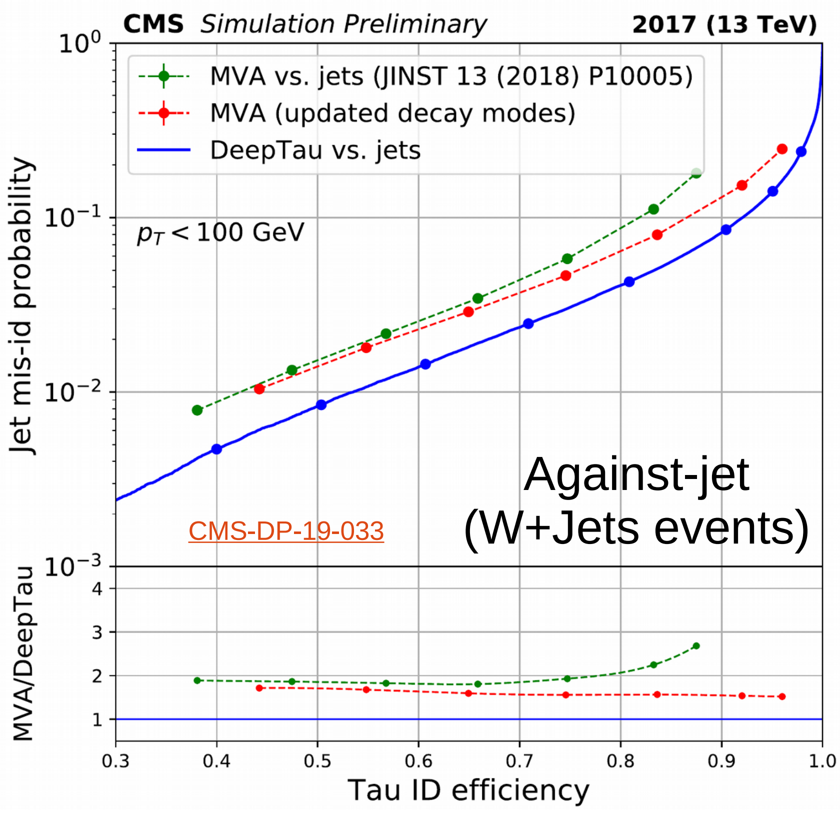


- The DeepTau against-muon (electron) classifier allows to obtain a noticeably higher rejection of $l \rightarrow \tau_h$ fakes compared to the cut-based (MVA-based) algorithms previously used

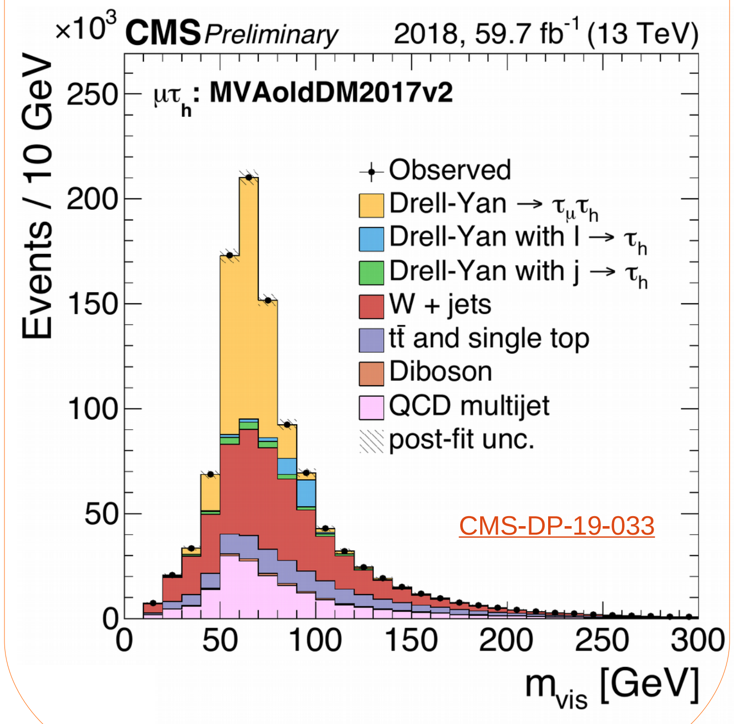
The against-jet classifier

- For the against-jet classifier, two separate studies are performed, the rejection of jets coming from tt-bar events, and of jets coming from W+Jets events





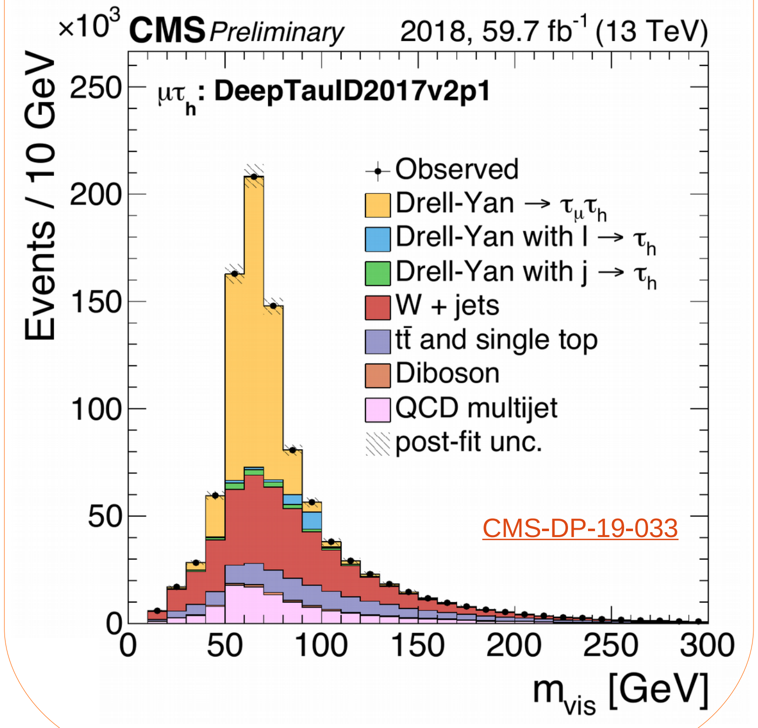
m_{vis} distribution using JINST 13 (2018) P10005



With DeepTau IDs the yield from **genuine τ_h** <u>increases by ≈ 20%</u>, while the yield from **fakes** <u>decreases by ≈ 23%</u>

Note: in both plots modelled contributions are fit to data

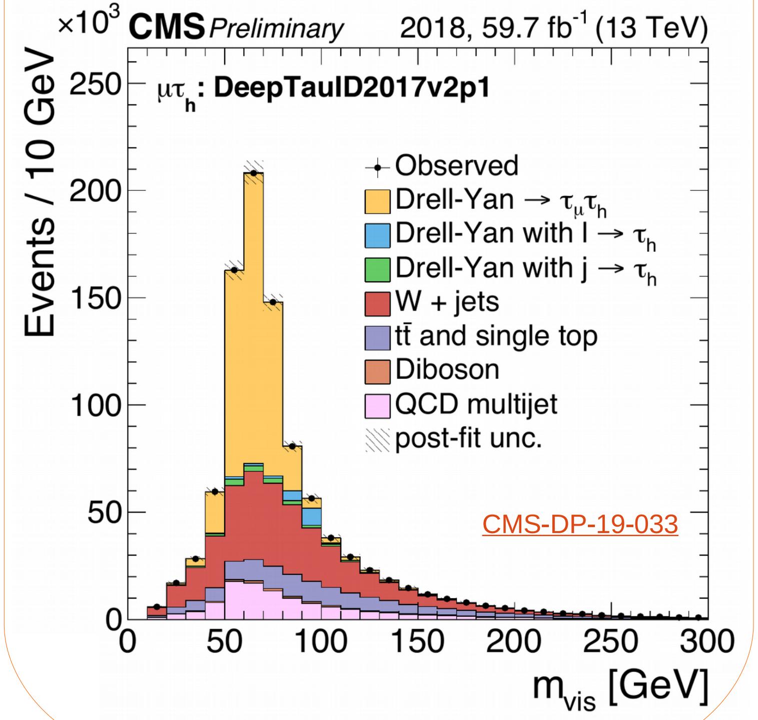
m_{vis} distribution using DeepTau IDs



Conclusions

- The tau identification exploiting deep learning techniques has
  - noticeably reduced the misidentification rate
  - increased the fraction of collected genuine hadronic taus
- Data/MC agreement has also improved, with correction SF now of the order of ~10%
- Other algorithms have been developed to study more specific kinematics:
  - "Performance of the low-pT tau identification algorithm" (CMS-DP-2020-039)

m_{vis} distribution using DeepTau IDs



Conclusions

- Several new analyses obtained a noticeable gain by using the DeepTau IDs
  - Overview shown tomorrow at 9.30 in Higgs session (link)

Higgs boson measurements in final states with taus at CMS

Look forward to new physics results with τ leptons.

Andrea Cardini

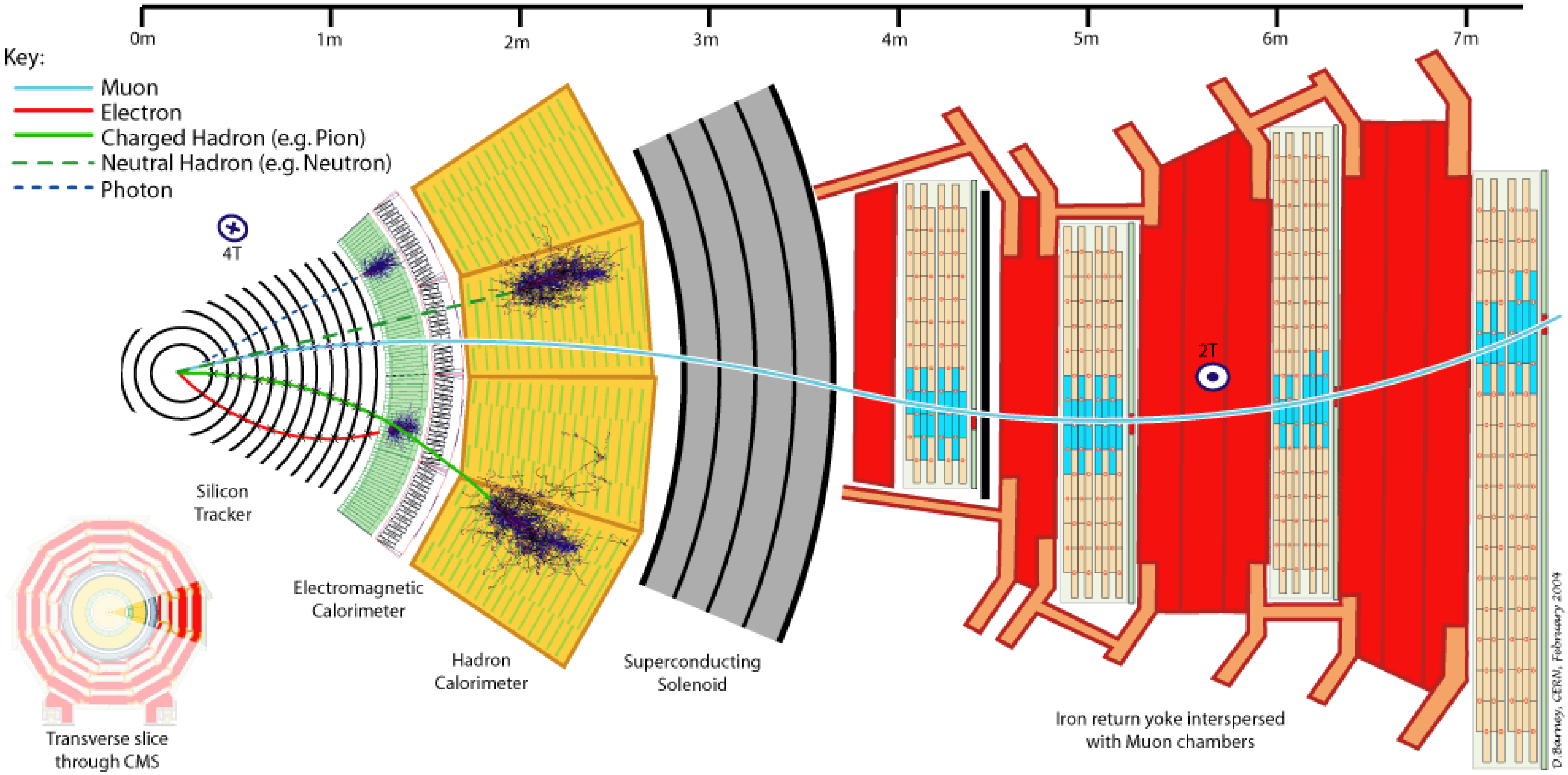
on behalf of the CMS Collaboration

ICHEP 2020 | PRAGUE

40th INTERNATIONAL CONFERENCE ON HIGH ENERGY PHYSICS

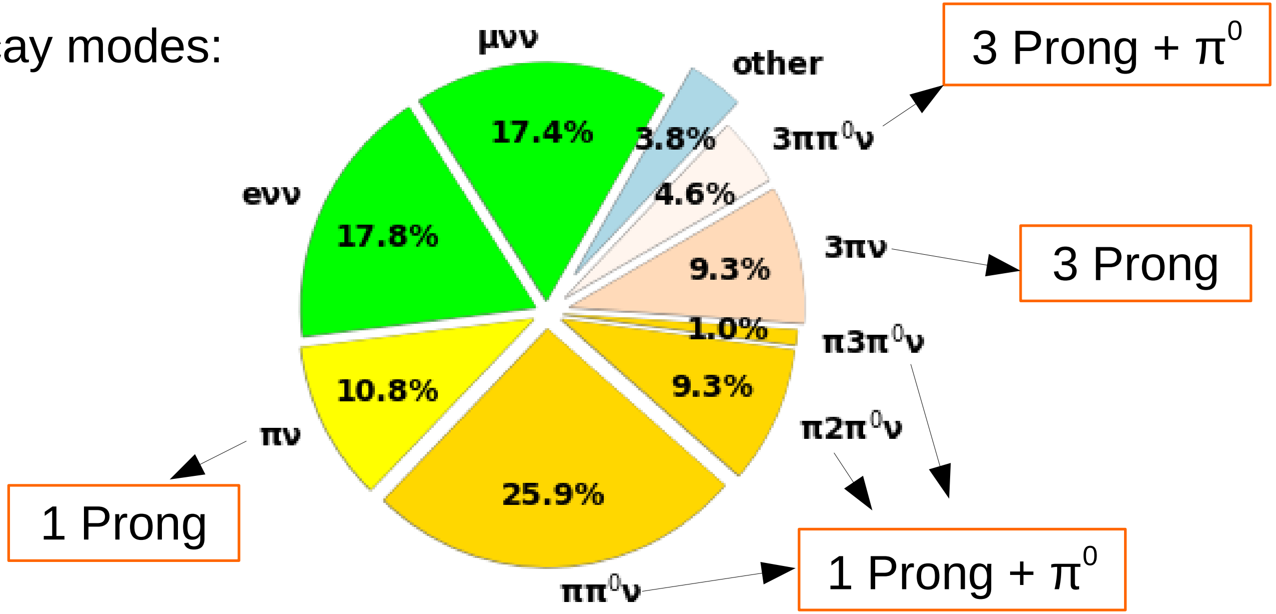
VIRTUAL CONFERENCE

CMS object reconstruction

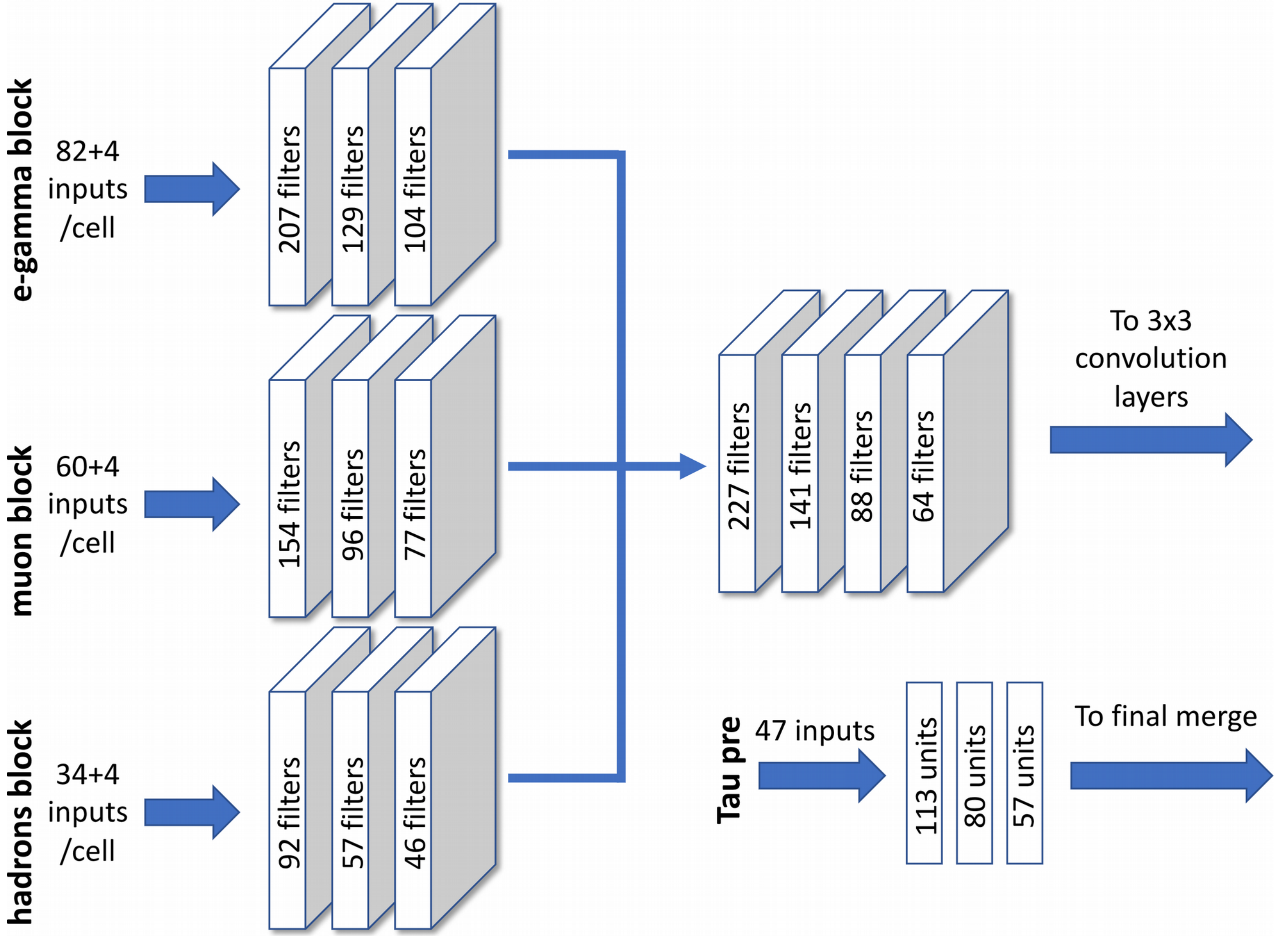


The Hadron-plus-strip algorithm

- The HPS algorithm combines:
  - PFCandindates of jets (p_T>0.5 GeV)
  - EM showers, i.e. PFCandidates which produce a deposit in the ECAL in a dinamic window size adjusted as a function of the e/γ p_T (minimum p_T > 1 GeV)
- It can reconstruct 4 decay modes:
  - 1 Prong
  - 1 Prong + π^0
  - 3 Prong
  - 3 Prong + π^0

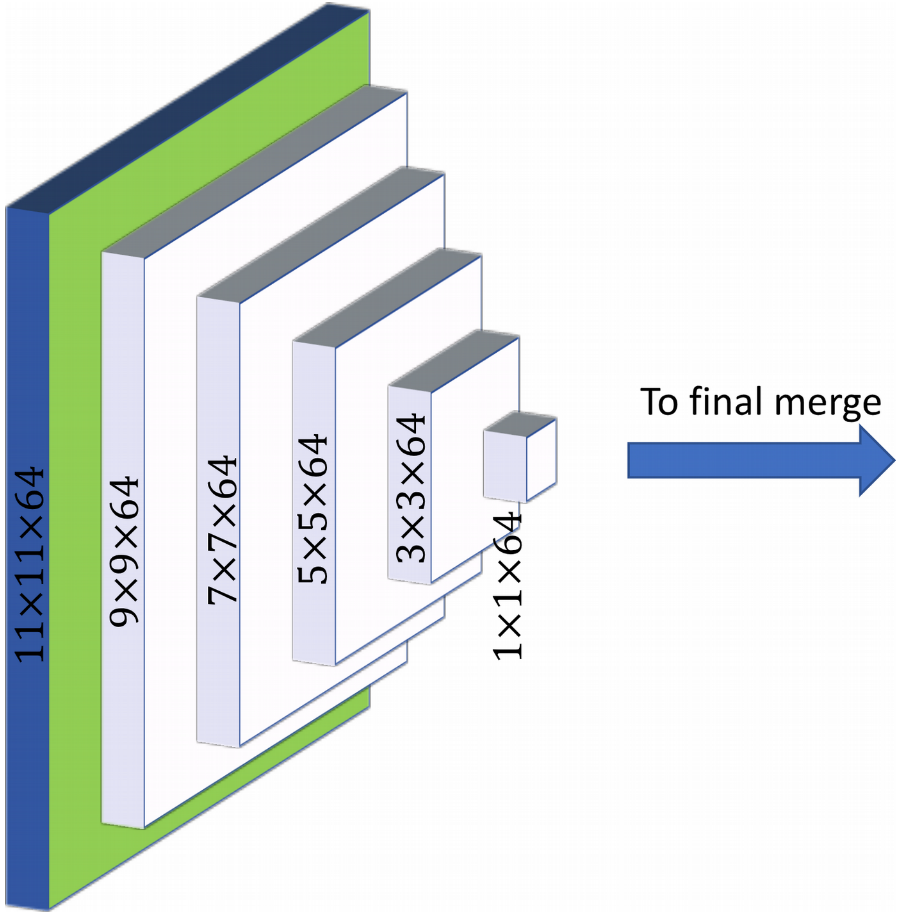


Pre-processing



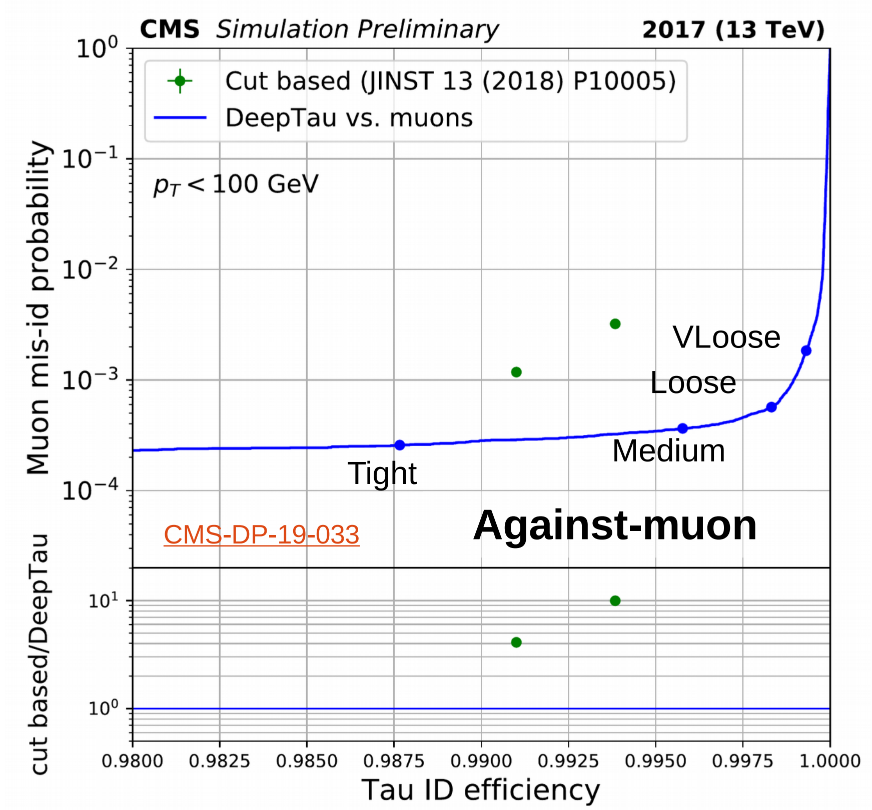
- To reduce the dimension of the problem the 188 features/cell are preprocessed through 3 convolutional layers with a window size of 1x1
- The outputs are concatenated and processed via 4 more layers
- The output of the pre-processing is a 11x11(21x21) x 64 array for the inner (outer) grid
- The high-level features are pre-processed by 3 dense layers

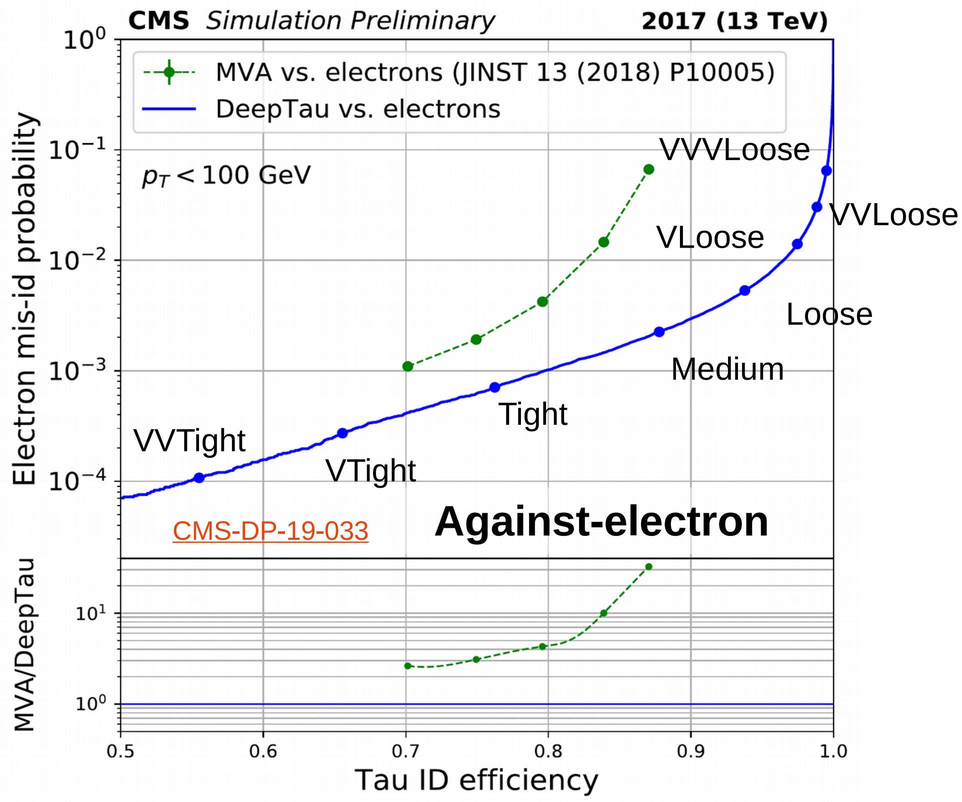
Convolution and final merge



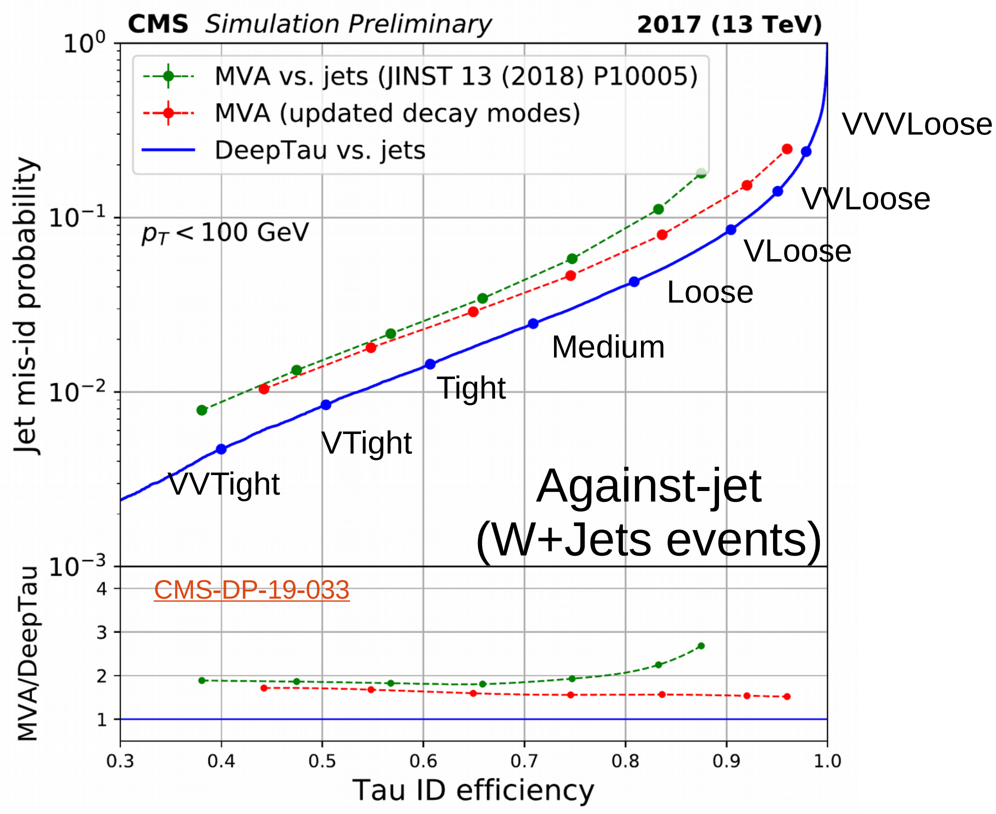
- After the pre-processing inner and outer features are processed through convolution layers with 3x3 window size
- The final layer flattens the grid into an array
- The number of layers used is:
  - 5 for inner features
  - 10 for outer features
- For the final merge the ouputs of the previous modules are concatenated and processed via 4 dense layers

Working points in CMS: against-lepton





Working points in CMS



- Several working points are defined on the classifiers
- This is done to allow analyses to combine selection on the different classifiers depending on the greatest sources of contamination
- The definition of the Working point is done on genuine τ_h with pT $\in$ [30,70] GeV using simulated $H \rightarrow \tau\tau$ events