Tau identification in CMS during LHC Run 2

Mohammad Hassanshahi

On behalf of the CMS collaboration

29 Sep 2021

TAU2021 conference

Introduction

- Several interesting physics studies are accessible at the LHC and, in particular, in tau final states. For example:
  - determining Higgs properties
  - searches for additional Higgs
  - lepton flavor universality (LFU) violation searches

- Instantaneous luminosity and energy increased in the Run II of the LHC (2015 - 18)
  - Challenge to reconstruct taus and identify them from fakes

- The CMS collaboration developed various machine learning techniques to improve reconstruction and identification of hadronic decays of taus:
  1. A complex convolutional DNN to identify taus from jets/muons/electrons (named "DeepTau").
  2. A BDT with XGBoost library to identify tau decay modes (named "MVA decay mode").
  3. An attention-based graph neural network called "ABCNet" to reconstruct low-pt taus.

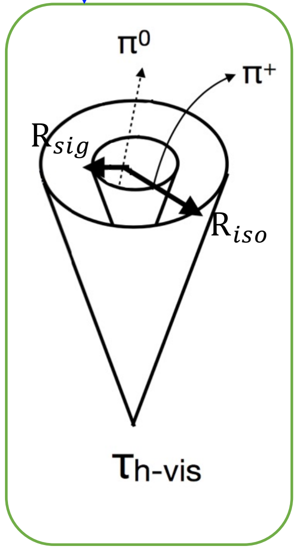
Tau reconstruction in CMS

- CMS employs particle flow (PF) algorithm to reconstruct individual physics objects.
  - μ, e, γ, charged (e.g. $\pi^{\pm}$, protons) and neutral (e.g. neutron) hadrons
  - It combines info. from all subdetectors.

- To reconstruct taus which decay hadronically, the *Hadron-Plus-Strip* (HPS) algorithm is used. There are 4 steps in this algorithm:

1. Anti-k_T jet-finding algorithm (with R=0.4) → reconstructed jets used as seeds for tau reconstruction. PF candidates within $R_{iso} = 0.5$ of the jet axis are considered for the next steps.
2. PF charged hadrons ($h^{\pm}$) are selected. e/γ within a dynamic-size "strip" in $\eta - \phi$ plane are collected and considered as π^0 candidates.
3. A decay mode is assigned to each τ candidate based on the num. of $h^{\pm}$ and π^0.
4. Extra conditions: mass and charge compatible with a τ decay. All $h^{\pm}$ and π^0 candidates must be within $R_{sig} = 3/p_T$ (GeV), with R_{sig} limited to 0.05-0.1 range. If more than one τ candidate selected → highest p_T kept

Decay mode	Resonance	$\mathcal{B}$ (%)
Leptonic decays		35.2
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$		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



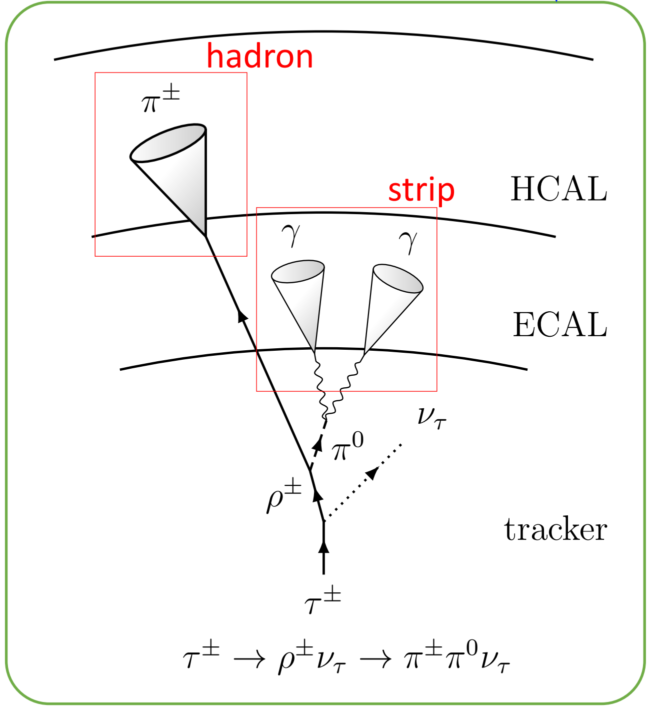


Image credit: Izaak Neutelings

Tau identification (*DeepTau*)

Identifying genuine taus

- Some objects can fake hadronic decay of taus:
  - **Electrons:** Can fake 1 charged-prong decays, in particular $\tau \rightarrow \pi\pi^0$ when an electron emits bremsstrahlung radiation
  - **Muons**: Usually fake $\tau \rightarrow \pi$ decay
  - **Jets**: collimated quark/gluon jets which pass mass/charge/… conditions of the HPS

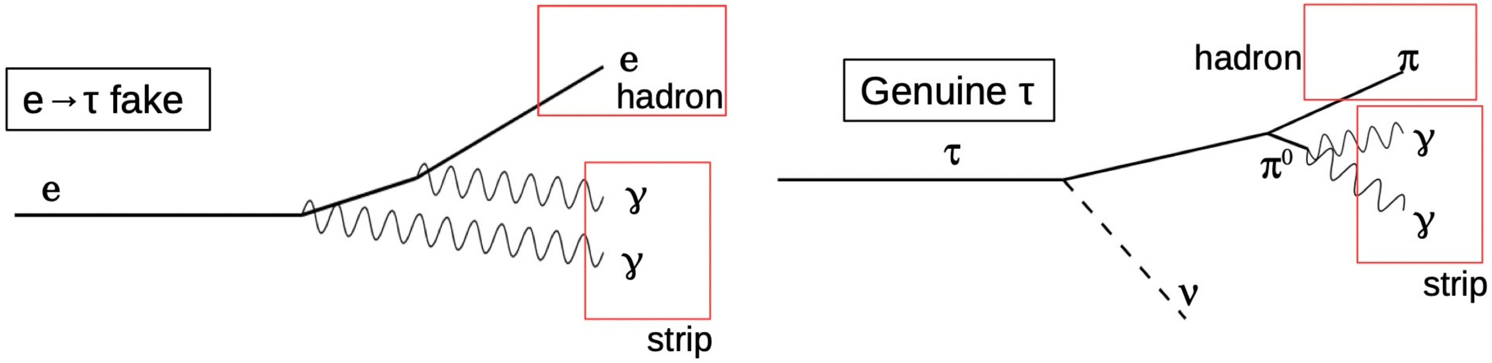


Image credit: Andrea Cardini

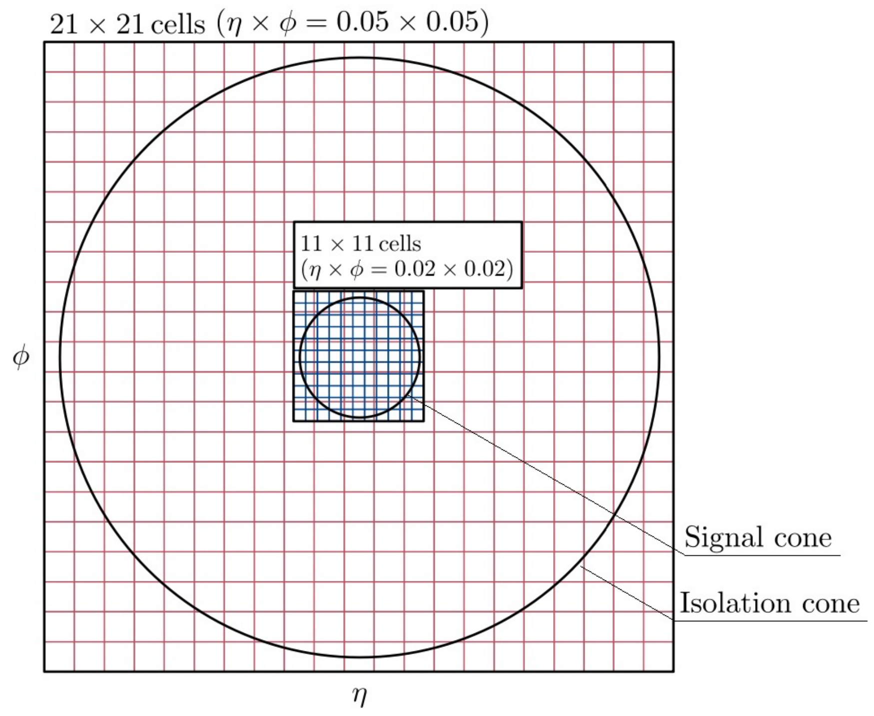
- We developed a deep convolutional neural network (***DeepTau***) to identify genuine taus from fakes.

Inputs of DeepTau algorithm

- The algorithm combines high-level and low-level features in a DNN.

- High-level features:
  - **General event** properties: e.g. estimated pile-up density
  - ***τ* candidate** properties: e.g. 4-momentum, #charged and #neutral particles, compatibility with Primary Vertex
  - 47 features

- Low-level features:
  - Targeting properties of ***τ* decay products**
  - Define two grids in $\eta - \phi$ space around τ axis
    - Inner grid: 11x11, finer, covering signal cone
    - Outer grid: 21x21, coarser, covering isolation cone
  - 188 features of particles (muon/electron/hadron) **per cell**



Architecture

- A total of ~100k input features used in the training.
- For **each cell**, low-level features are preprocessed independently for **e/ γ**, **muons**, and **hadrons.**
- The grids are reduced with CNN to 1x1 grids and then concatenated with high-level variables into the final DNN.
- Final output includes 4 classes: electrons, muons, hadronic taus, jets.
- The training is performed using NAdam algorithm.
- Loss function includes regular cross-entropy term plus binary focal loss term for tau against all backgrounds.

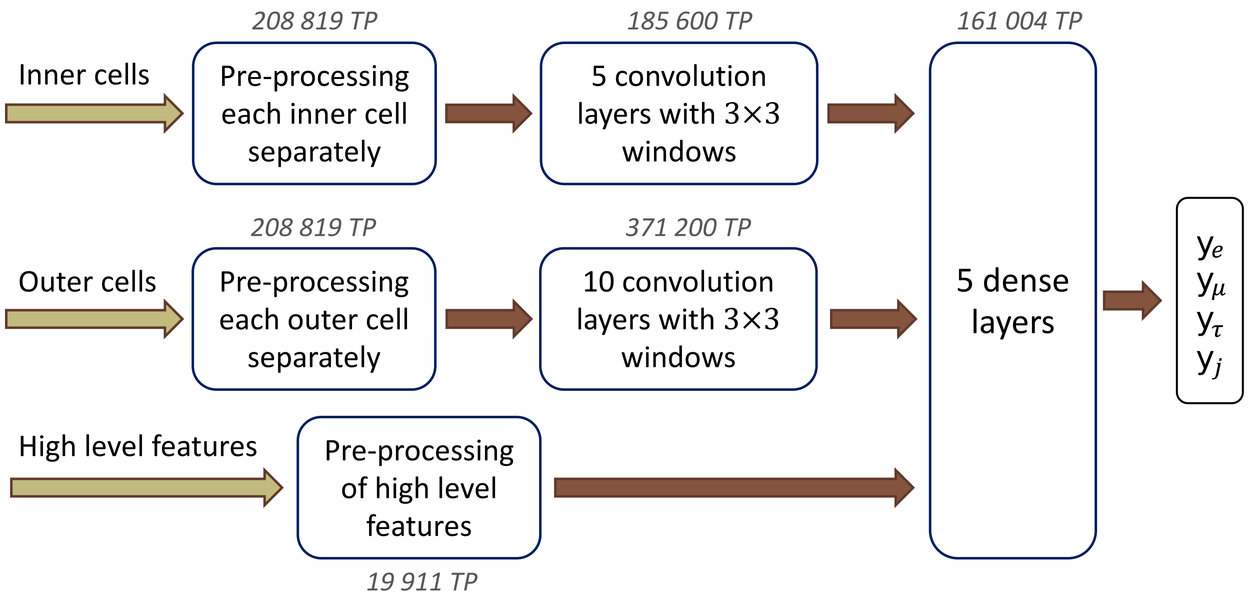
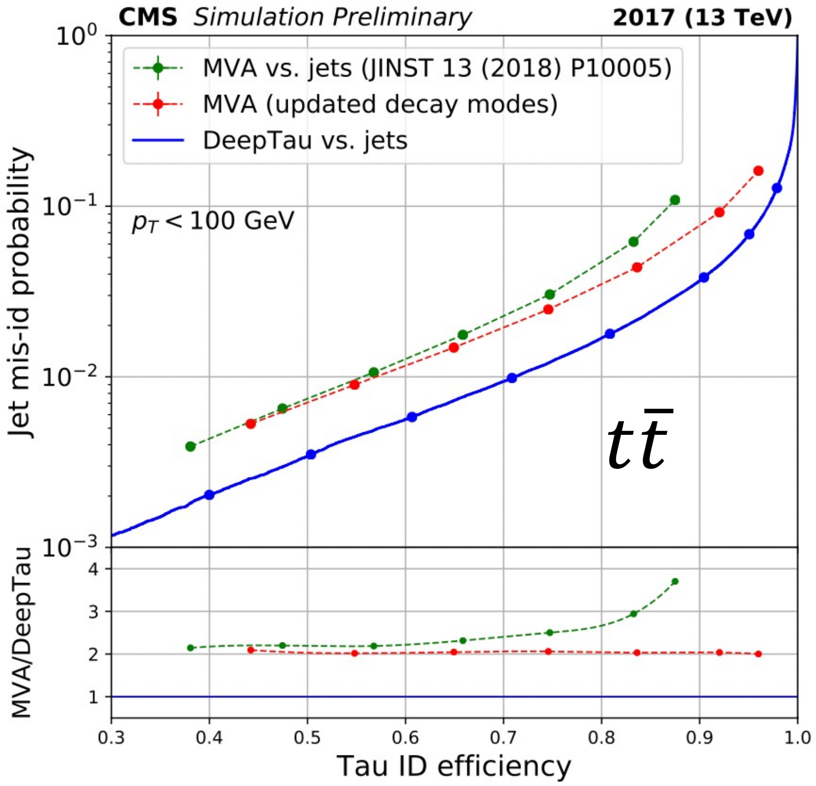


image source

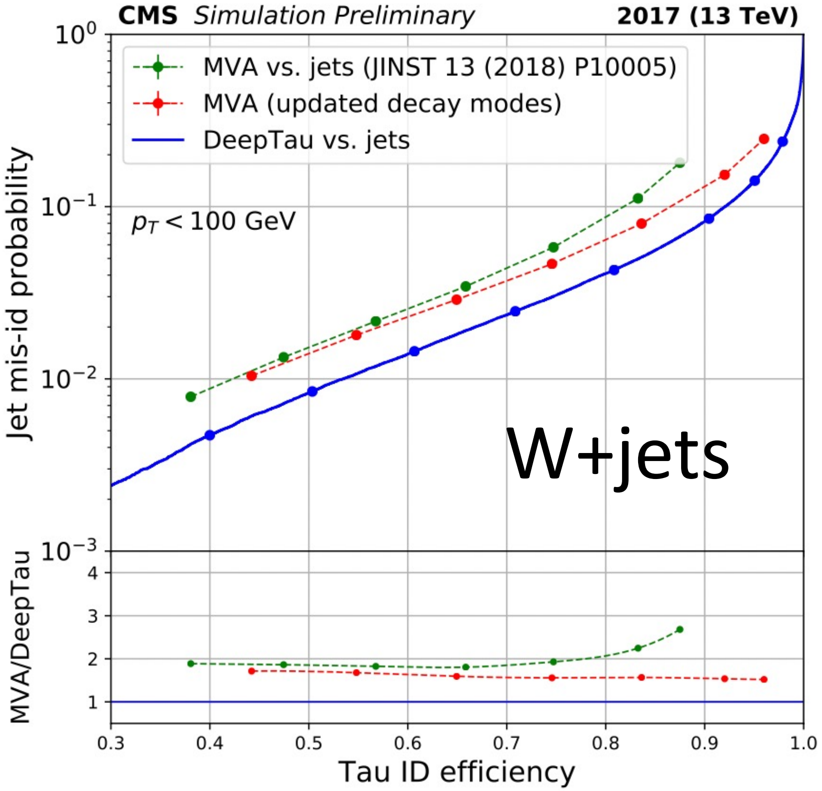
Final discriminator: $D_\alpha(\boldsymbol{y}) = \dfrac{y_\tau}{y_\tau + y_\alpha}$

Discrimination against jets

- The performance of DeepTau against jets is checked using $t\bar{t}$ and W+jets simulated events. DeepTau significantly outperforms the previous algorithm in all working points.

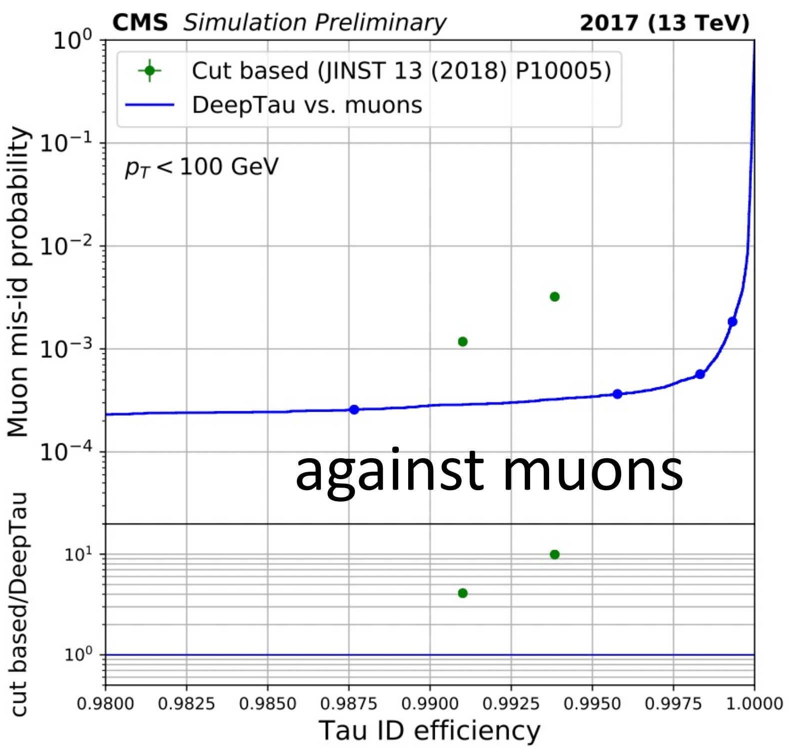


CMS-DP-2019-033

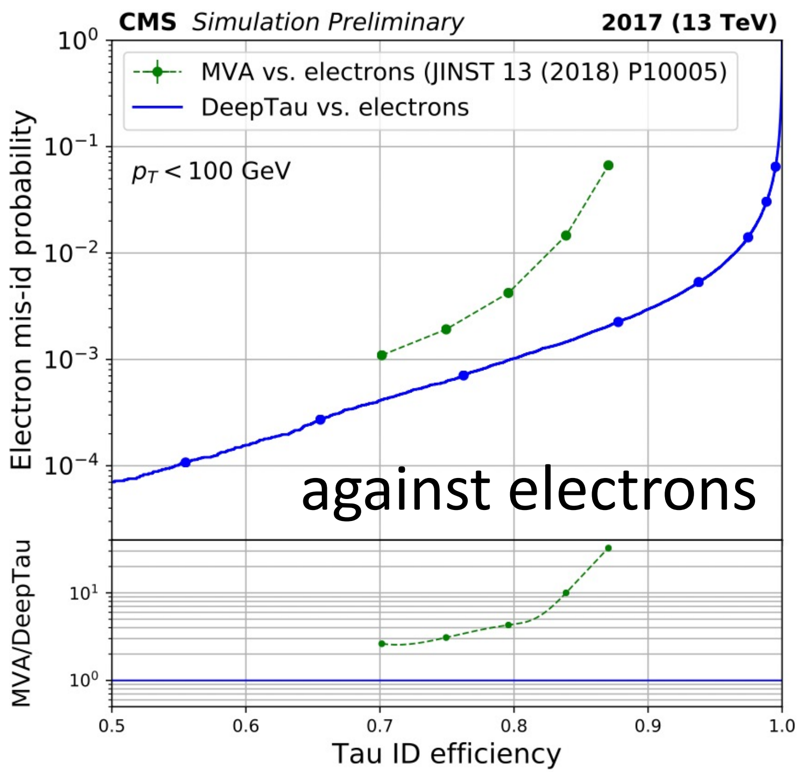


Discrimination against electrons and muons

- Likewise, DeepTau noticeably enhances the discrimination of taus against muons and electrons.



CMS-DP-2019-033



❖DeepTau has already been used in several CMS analyses, e.g. Higgs CP, Heavy Higgs search, LFV in Higgs decay and Higgs production cross section.

Decay mode identification (*MVA decay mode*)

Motivation

- The Hadron-Plus-Strip algorithm is developed to reconstruct taus. But it is not *optimized* to distinguish τ decay modes.
  - The downside is that for e.g. all τ decays to one π and multiple π^0's are reconstructed as $\tau \rightarrow \pi\pi^0$

- There are important physics measurements of which sensitivity differ by τ decay modes: e.g. $H \rightarrow \tau^+\tau^-$ CP analysis (see A. Cardini's talk today), and τ polarization measurement in Z decay.

- To improve decay mode identification, two independent classifiers were trained for 1 and 3 charged-prong decays. The reconstructed decay modes are named as *MVA decay mode*.

- The classifiers are based on boosted decision tree (BDT) algorithm from XGBoost library.

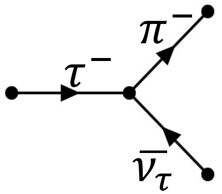
Classifiers

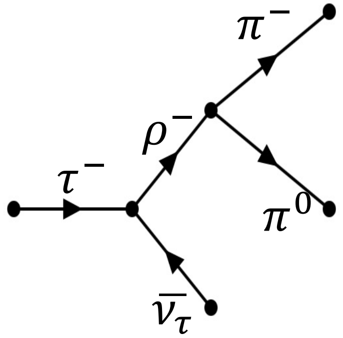
- Four outputs (final states) for the 1 charged-prong decay: π, $\pi\pi^0$, $\pi 2\pi^0$, and "other" category.

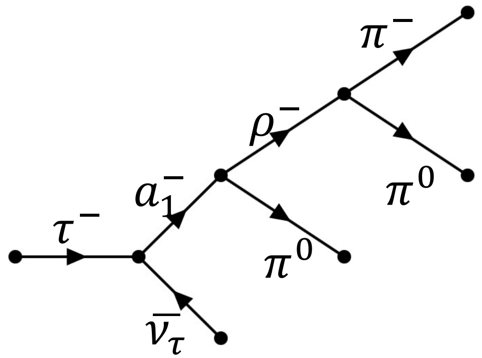
- The decays differ by the number of π^0's in the final state and their intermediate resonances. In this classifier, we exploited these information, along with the angular and kinematic quantities of the reconstructed decay products.

- The main features used in the training:
  - invariant mass of π^0 † and of ρ
  - kinematics of τ_{vis} and the decay products
  - angular variables between π^0 constituents and between (π, π^0) pair
  - HPS decay mode

- Similar strategy for the 3 charged-prong decays, but with three outputs: 3π, $\pi^0 3\pi$, and "other" category.







†. π^0 4-momentum is the same as strip with slight modification (e.g. low energy photons removed)

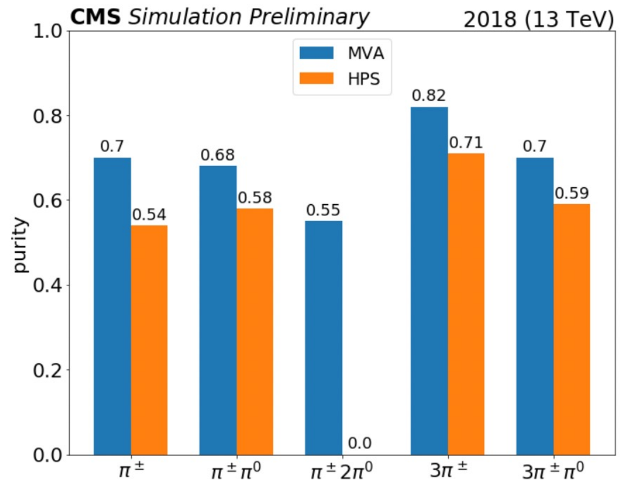
Efficiency and Purity

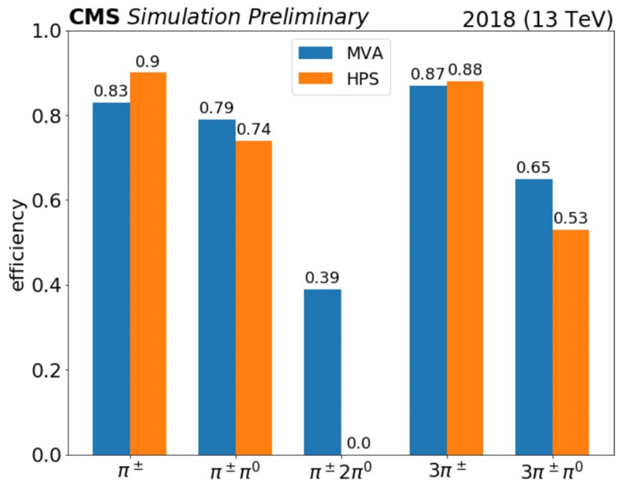
- The performance of decay mode reconstruction with *MVA decay mode* versus the default HPS algorithm is compared.

- The purity has improved by 10-55%, depending on the exact decay mode.

- The efficiency of the decay modes with at least one π^0 has increased by 5-40%.

- In the $H \rightarrow \tau^+\tau^-$ CP analysis with the full Run 2 CMS data set, using the *MVA decay mode* enhanced the CP sensitivity by $\approx 20\%$.

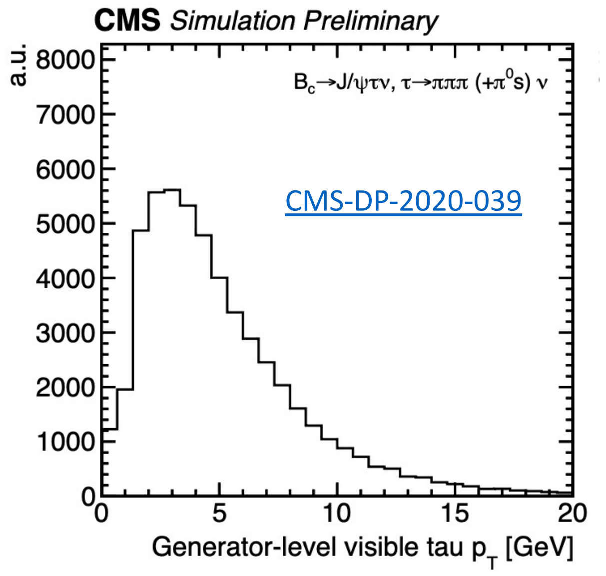


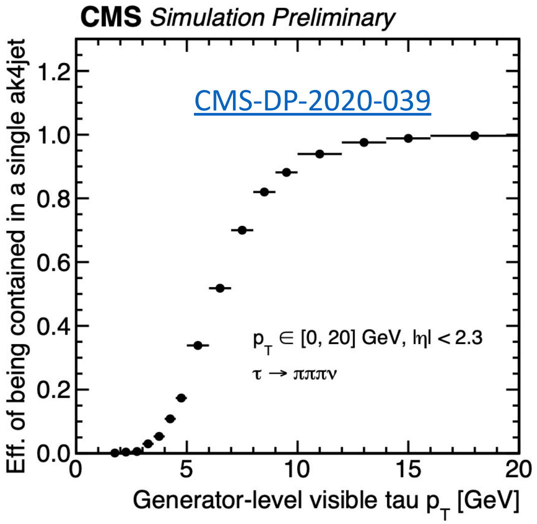


[CMS-DP-2020-041]

Low-p_T τ identification

- Anomalies in LFU and g-2 → growing interest for cross-checks in other experiments.
- For such analyses (and even more, e.g. compressed SUSY searches) an efficient reconstruction of low-p_T (p_T<15 GeV) taus is required.
- The HPS algorithm is not acceptable in this regime as the τ decay products may not be contained in a jet seed
  - high magnetic field + low p_T charged hadron → large spread over $\eta - \phi$ plane
- We developed a dedicated algorithm to identify low-p_T 3 charged-prong decays of taus (3π and $\pi^0 3\pi$) **without** using seeding jets.
- The algorithm is trained on and optimized for taus from B meson decays. But it can be extended for other physics studies.



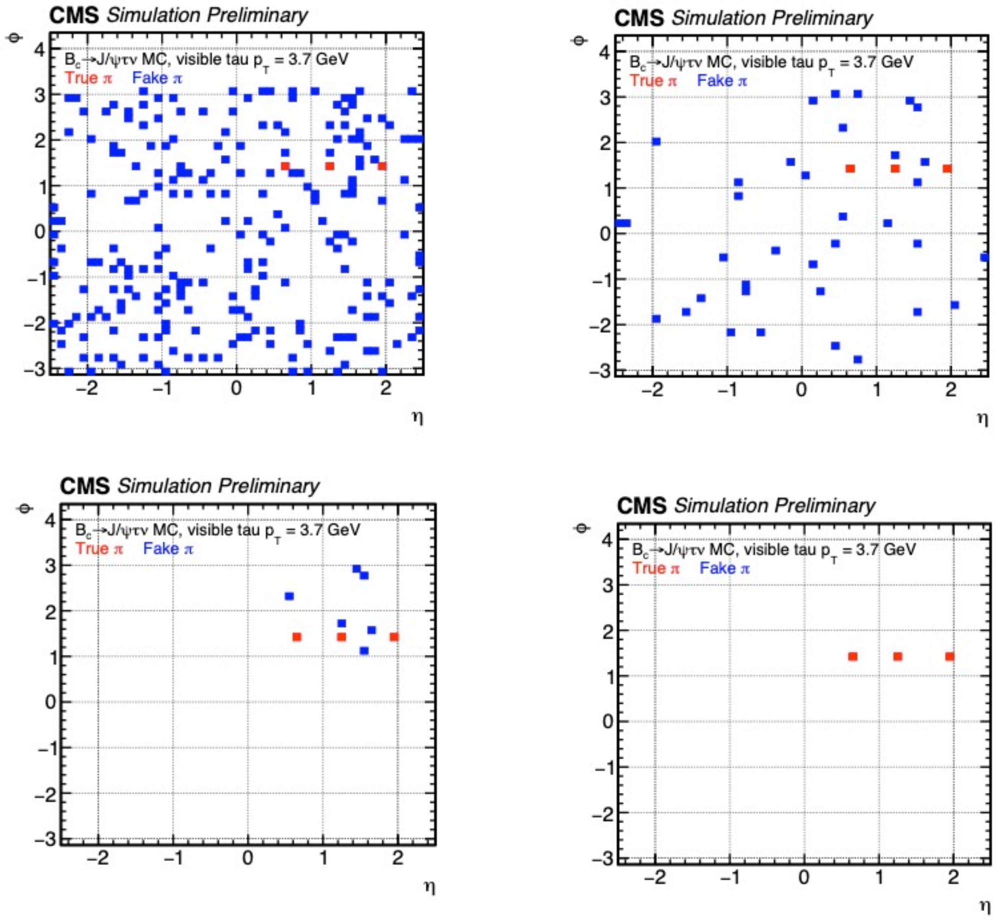


Reconstruction algorithm

- The τ reconstruction algorithm is as follows:
  - All PF charged pions are collected
  - Keep only pions in the vicinity of the Primary Vertex (PV), from which B_c is produced.
  - *ABCNet*: An attention-based graph neural network → assigns a prob. to each $\pi^{\pm}$ for originating from a tau → fakes suppressed by requiring ABCNet score > 0.1443
    - Input features to DNN are 4-momentum, charge, and distance from PV.
  - Choose the highest-p_T τ candidate satisfying specific criteria: conditions on sum of ABCNet score, τ flight length significance and τ vertex prob.

- ABCNet has two main advantages:
  - Graph NN → Byproducts of p-p collisions are treated similar way as they are recorded in the detector.
  - Attention based → local features exploited→ more efficient architecture

CMS-DP-2020-039



Identification efficiency

- The eff. of identifying the right combination of $3\pi^{\pm}$ in the $B_c \to J/\psi\tau\nu$ decay is shown as a function of generated tau p_T.

- Eff. is defined as:

$$\varepsilon = \frac{\text{Tau is reconstructed and the three charged pions are the right combination}}{\text{All events with 3-prong tau at the generator-level}}$$

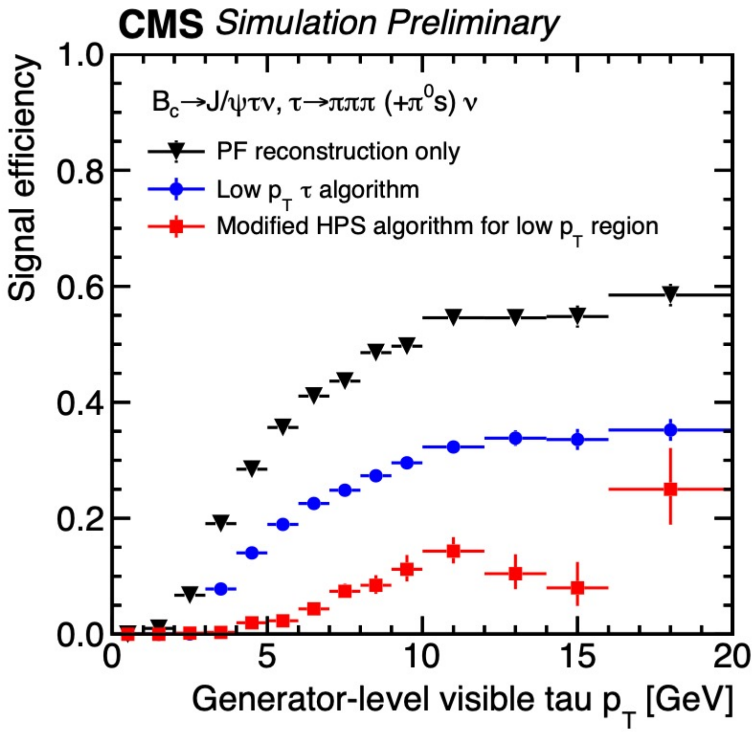
- The "modified" HPS algorithm (red) is the HPS but with twice larger distance parameter for jet seeding $(0.4 \to 0.8)$.

- Low-p_T algorithm in blue.

- The black dots is the reco. eff. of all three PF charged pions (max. eff. one can get with the existing PF algorithm)

- The algorithm will soon be used in $R(J/\psi)$ analysis. Can potentially be extended to tau g-2, $B_s \to \tau\tau$, $R(D^*)$, etc.

Sum(ABCNet) > 2.3 applied.



CMS-DP-2020-039

Conclusion and outlook

Conclusion and outlook

- Various interesting physics studies at the LHC include taus in the final state.

- The CMS collaboration has developed advanced ML techniques in order to improve τ identification. These techniques have already been (and will be) used in many physics analyses for the data taken during the Run-2 of the LHC.

Algorithm	ML architecture	(will be) used in
DeepTau (tau Id. against jets/muons/electrons)	Deep CNN	Higgs properties, BSM searches, etc.
MVA decay mode (tau decay mode Id.)	BDT with XGBoost	Higgs CP, tau polarization in Z decay
Low-p_T tau reconstruction	Attention-based graph NN	$R(J/\psi)$. Potentially in tau g-2, compressed SUSY, etc.

- Motivated by new-physics searches with tau final states, CMS will continue improving τ identification with state-of-the-art computing techniques.

Backup

Strip clustering in the HPS algorithm

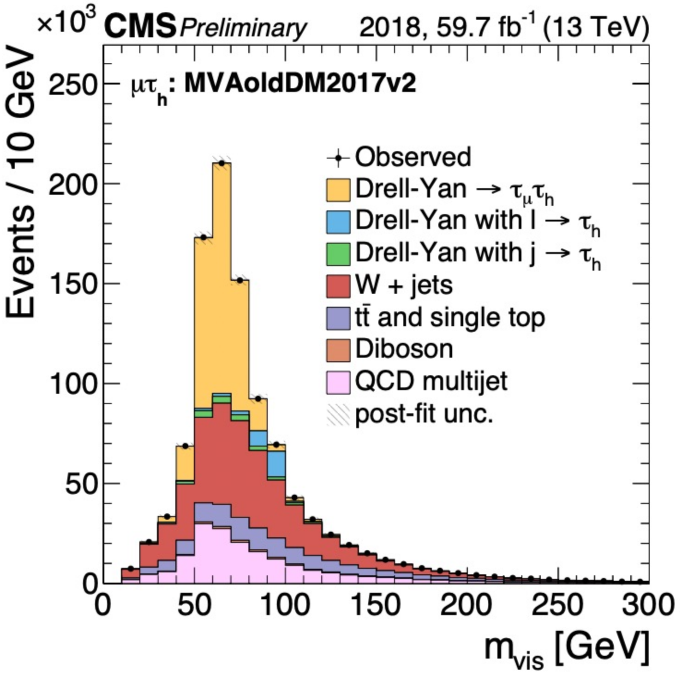
- The strip clustering in the HPS algorithm:
  1. The highest p_T e/γ not yet included in any strip seeds a new strip. Initial position set to the (η, ϕ) of the seed.
  2. The 2nd highest p_T e/γ within a cone size of $\Delta\eta = f(p_T^{e/\gamma}) + f(p_T^{strip})$ and $\Delta\phi = g(p_T^{e/\gamma}) + g(p_T^{strip})$ is merged into the strip. f and g are defined as:
     - $f(p_T) = 0.20\, p_T^{-0.66}$
     - $g(p_T) = 0.35\, p_T^{-0.71}$

     (Our simulations show that the above functions result in covering 95% of e/γ from a tau decay)
  3. The strip position is recomputed by the p_T-weighted average of all e/γ in the strip.
  4. When no further e/γ is found within the $\Delta\eta \times \Delta\Phi$ window, the strip construction ends. A new strip will be constructed using the highest-p_T candidate not yet contained within a strip.

- More details [JINST 13 (2018) P10005].

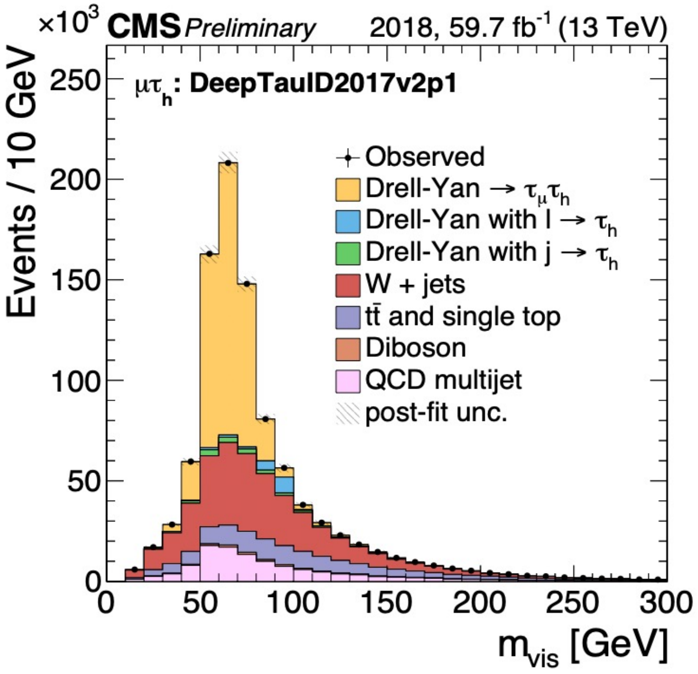
DeepTau improves purity

- The following plots show how using DeepTau increases the purity of $Z \to \tau\tau \to \tau_h\tau_\mu$ events. In both plots, contributions are fit to the data.



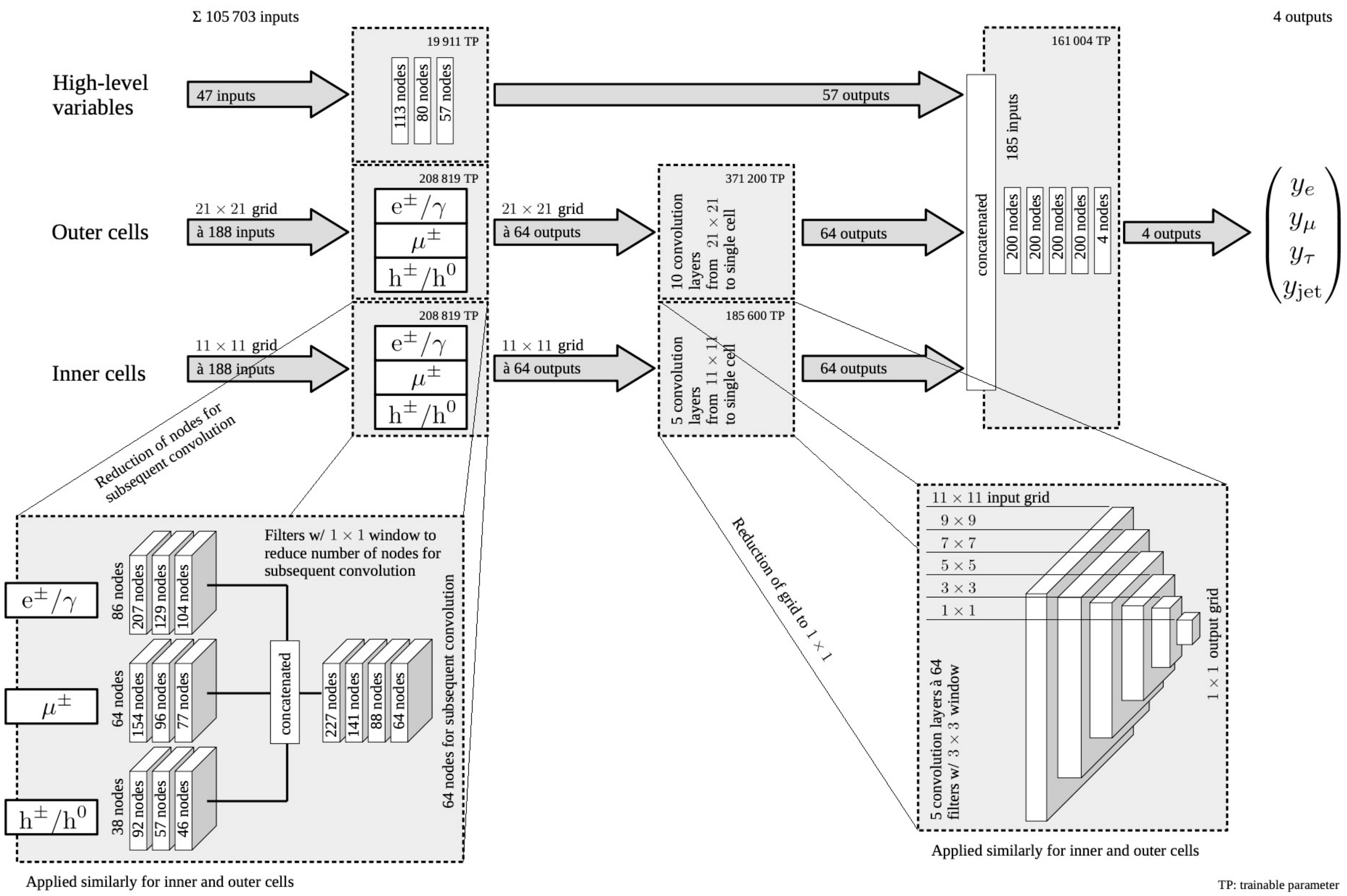
DeepTau increases genuine tau by ≈ 20% and decreases fakes by ≈ 23%

CMS-DP-2019-033

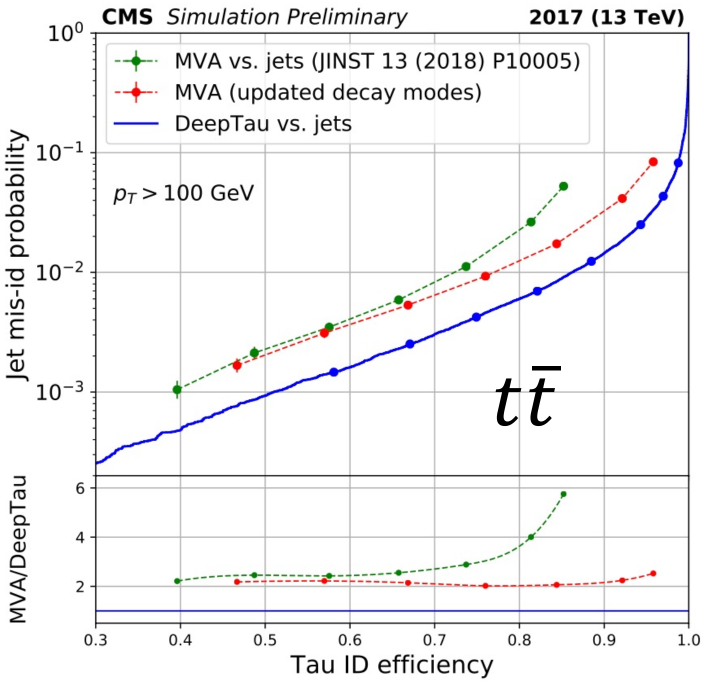


- DeepTau has already been used in several CMS analyses, e.g. Higgs CP, Heavy Higgs search, LFV in Higgs decay and Higgs production cross section.

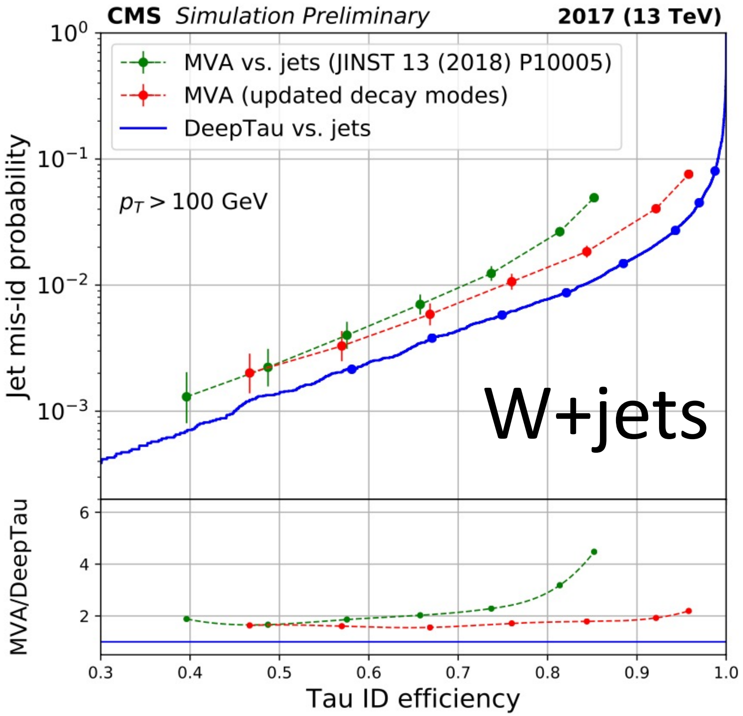
DeepTau full architecture



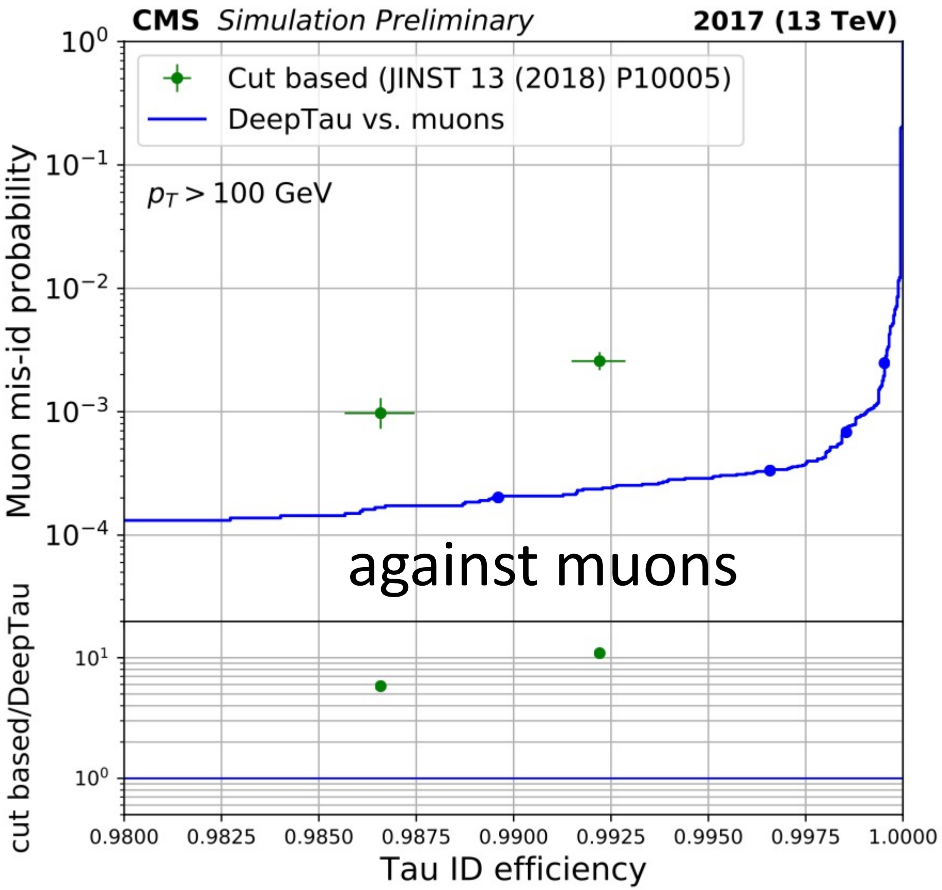
DeepTau against jets (pt>100)



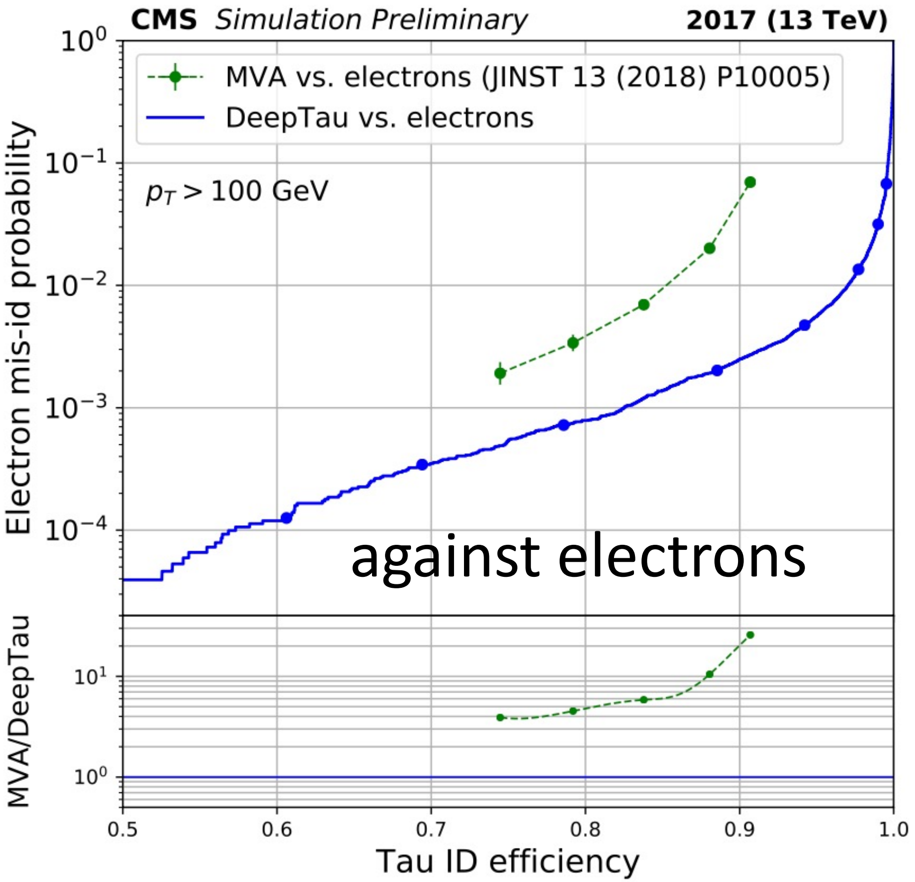
CMS-DP-2019-033



DeepTau against muons/electrons (pt>100)



CMS-DP-2019-033

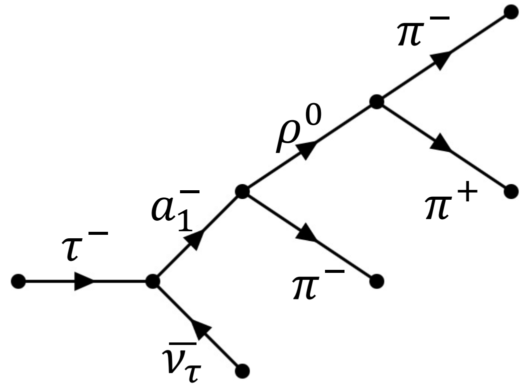


3 charged-prong decays classifier

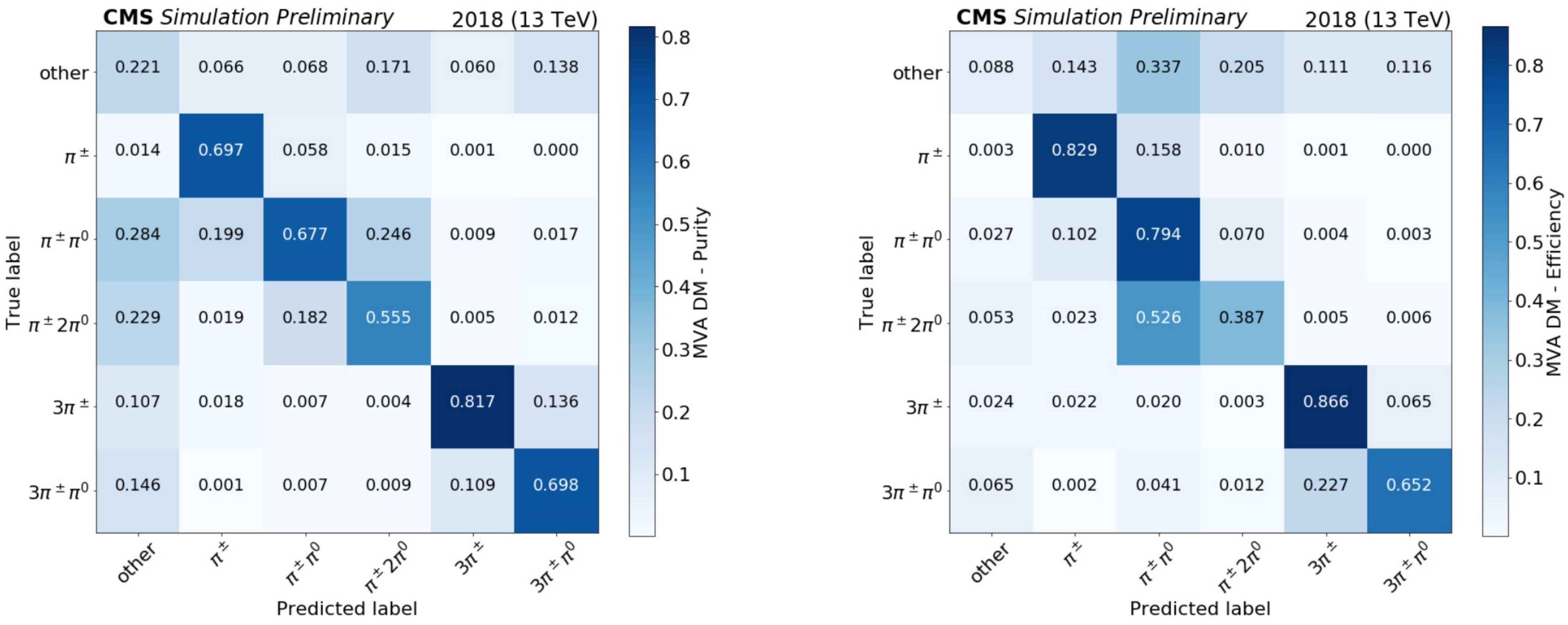
- Three outputs: 3π, $\pi^0 3\pi$, and "other" category.

- The strategy is similar to the 1 charged-prong case. Again the decays differ in the intermediate resonance and the number of π^0 in the final state.

- The features are also similar to the 1 charged-prong case. Additionally, for the pairs of charged pions, kinematic and angular quantities are added..



MVA decay mode confusion matrix



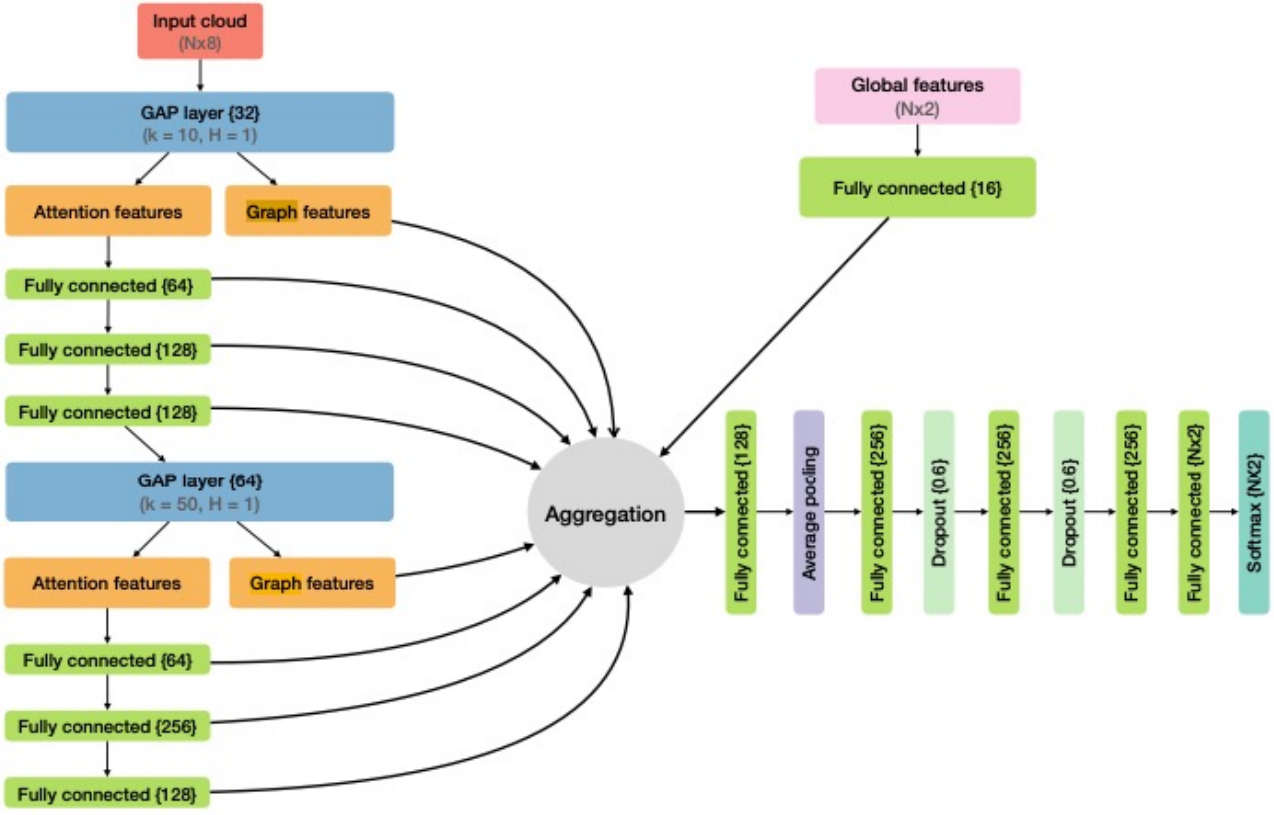
The confusion matrices of the MVA DM normalized by column (left) and by row (right).

CMS-DP-2020-041

ABCNet

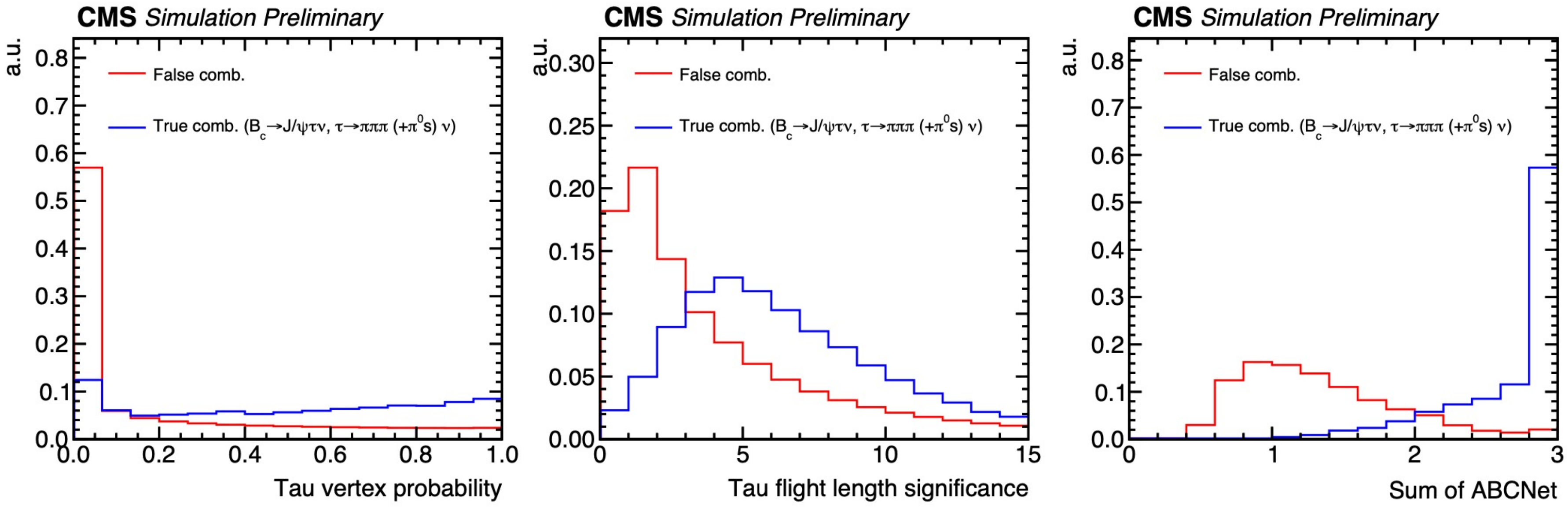
- ABCNet paper: [EPJ. Plus 135 (2020) 463]

- The architecture used in the low-p_T algorithm is shown on the right.



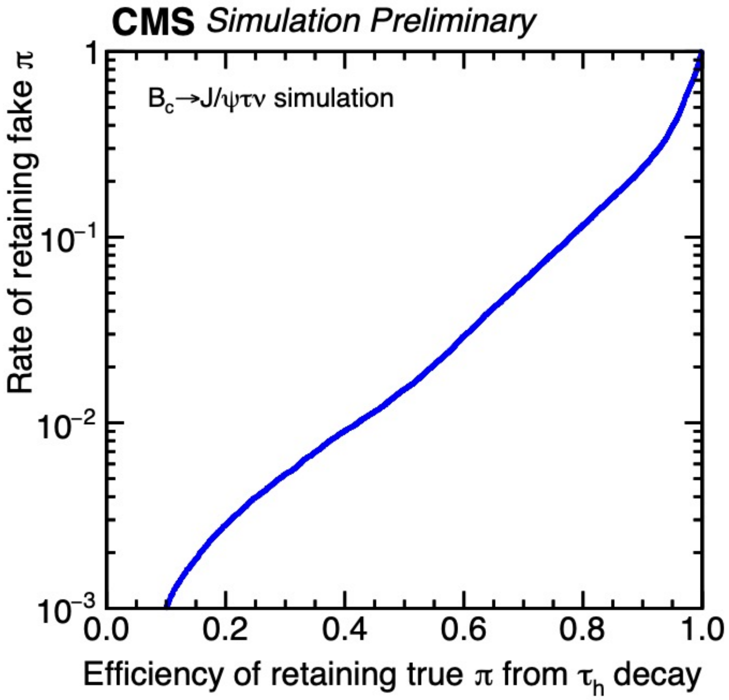
$B_c \to J/\psi\tau\nu$ example

- The final selection criteria of three charged pions can differ based on the efficiency and purity desired for an analysis. In the $B_c \to J/\psi\tau\nu$ example, the highest p_T candidate with the following criteria is selected:
  - τ vertex prob $> 10\%$
  - τ flight length significance $> 3\sigma$
  - Sum of ABCNet scores > 2.3



[CMS-DP-2020-039]

low pT: ABCNet ROC curve



CMS-DP-2020-039

Fig4: The ROC curve of the ABCNet pre-filtering for a single pion. The x-axis denotes the efficiency of retaining the true pion from the τ_h decay. The y-axis shows the rate of retaining a fake pion. The ABCNet is trained using τs from $B_c \to J/\Psi\tau\nu$ decays.

Low pt: signal eff. vs pile-up

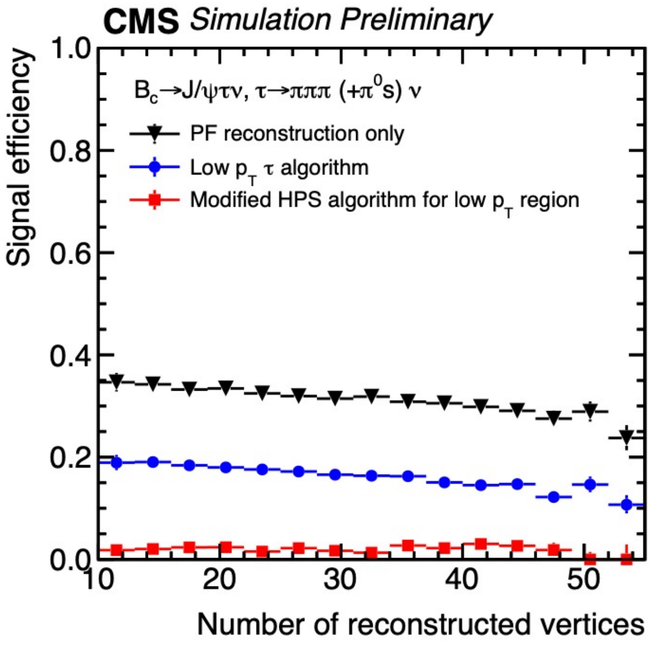


Fig9: The signal efficiency as a function of the number of reconstructed vertices, using a fixed cut of sum(ABCNet) > 2.3. The same legend applies as in Fig.8.

CMS-DP-2020-039