PHYSICS AT THE TERA SCALE
Helmholtz Alliance

CMS Compact Muon Solenoid

Identification of hadronically decaying tau leptons

28.11.2017

Speaker:
Olena Hlushchenko

Physics Institute III B
RWTH AACHEN UNIVERSITY

Introduction

RWTH AACHEN UNIVERSITY

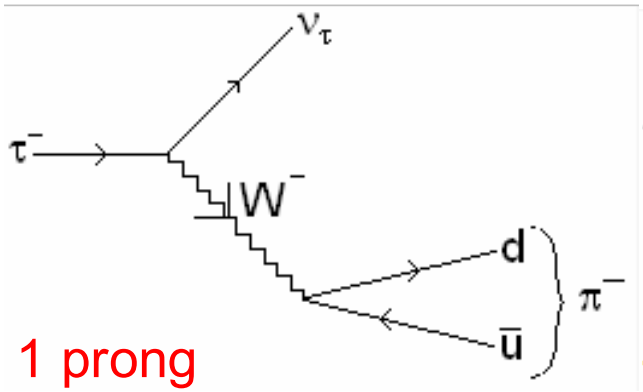
Motivation for measurements:

- "Beyond the Standard Model" **new physics searches**:
  - → $Z' \to \tau\tau$, $W' \to \tau\nu$
  - → SUSY searches with stau's
  - → Lepton flavor violation, as $Z \to \tau l$
- Higgs boson Yukawa couplings and CP studies
- Polarization studies in $Z \to \tau\tau$
- Any other analysis involving tau leptons

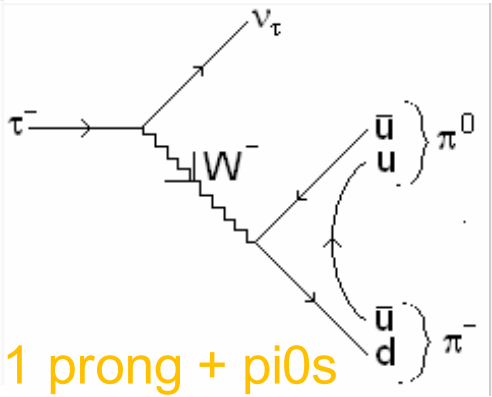
Tau lepton hadronic decay modes

RWTH AACHEN UNIVERSITY

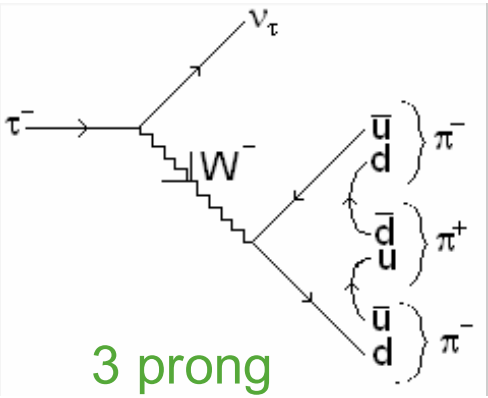
- It is the only lepton heavy enough (1.777 GeV) to decay into hadrons
- tau decays always involve neutrinos
- τ_h- tau hadronic modes, $\mathcal{B}r(\tau \to \tau_h)$ = 64.8%
- **The main hadronic decay modes are grouped by the numbers of charged (prongs) and neutral hadrons**



1 prong



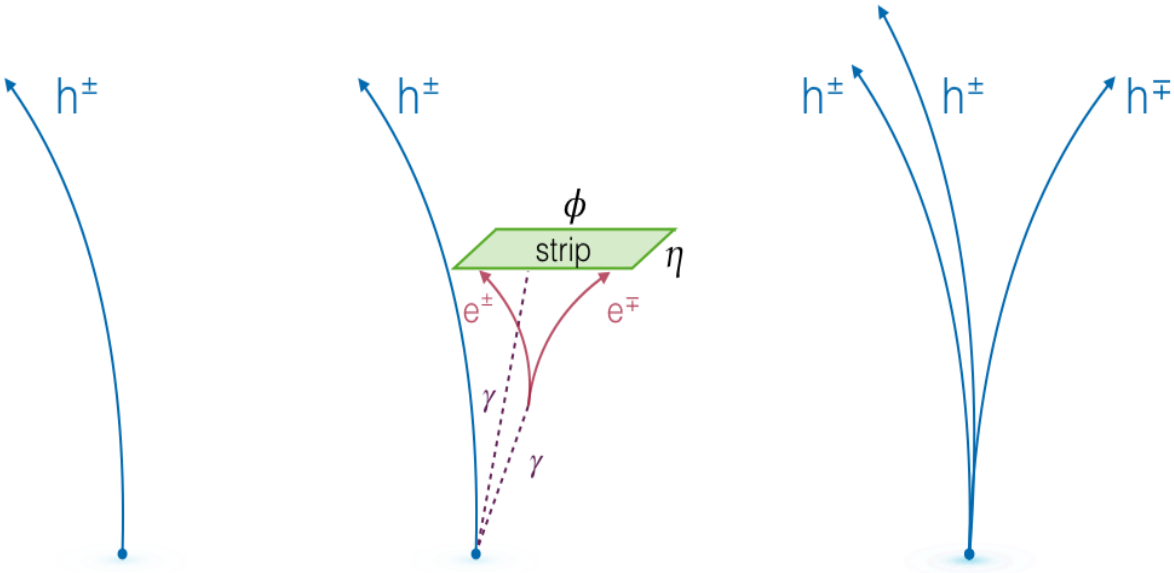
1 prong + pi0s



3 prong

Reconstruction of τ jets

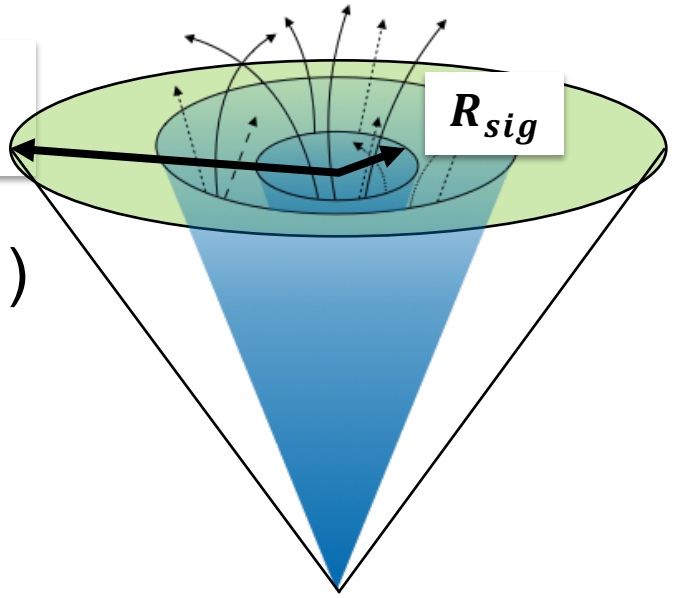
- A ”hadrons-plus-strips” (**HPS**) algorithm is run on top of the reconstructed particle flow (PF) jets:
  - → Reassembles **jets** to its constituents (photons, muons, hadrons)
  - → Reconstructs the π^0 from the electrons and photons
  - → Reconstructs possible **decay modes (DM)**
  - → **The DM with highest transverse momenta and charge/strip multiplicity that pass the mass cut-window is assigned to each jet**



QCD jets rejection

RWTH AACHEN UNIVERSITY

- **Multijet production is a background** for every analysis with hadronically decaying tau's
- QCD jets can be separated from the genuine tau hadronic decays by the larger hadronic activity in a predefined 'isolation cone'
- **Tools to discriminate:** cut- and MVA-based isolation discriminators
- R_{sig} - cone of the tau jet (0.05 - 0.1)



Isolation

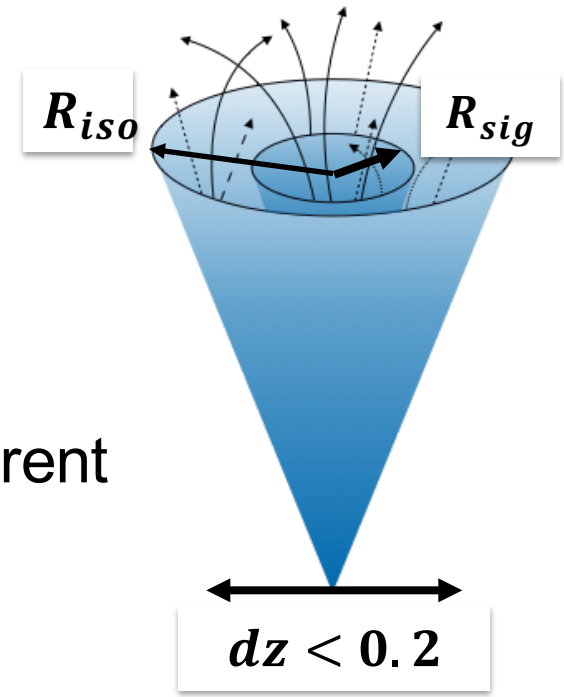
RWTH AACHEN UNIVERSITY

Sum runs inside of a so-called "isolation" cone

$$I_\tau = \sum p_T^{\text{charged}}(d_Z < 0.2\,\text{cm}) + \max\left(0, \sum p_T^{\gamma} - \Delta\beta \sum p_T^{\text{charged}}(d_Z > 0.2\,\text{cm})\right)$$

$$p_T^{\text{strip, outer}} = \sum p_T^{e/\gamma}(\Delta R > R_{\text{sig}}) < 0.10 \cdot p_T^{\tau}$$

- Isolation is sensitive to the PU
- MVA isolation on 23 input variables
  - → Tool: TMVA BDTG
  - → Individual terms of cut-based isolation are included into input variables
- Different isolation cones and DM configurations are required for different analyses

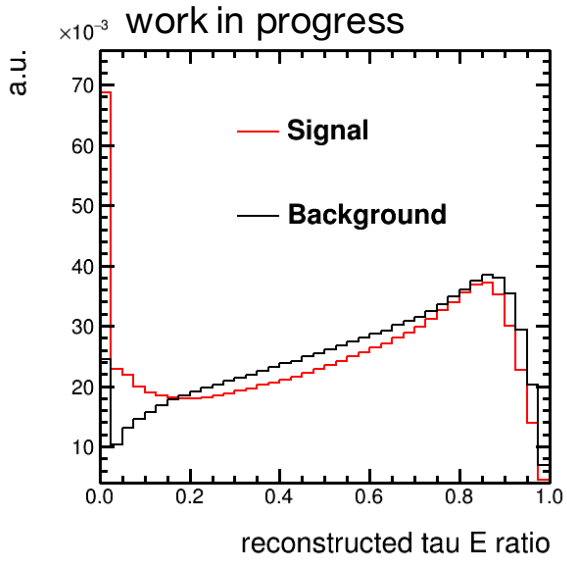


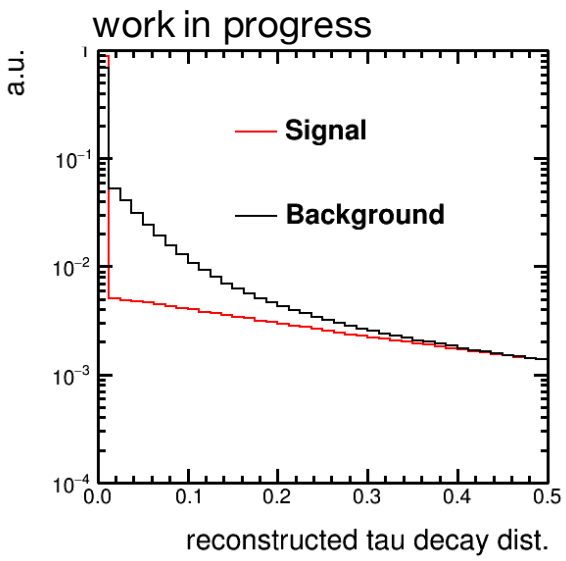
Classes of discriminating variables

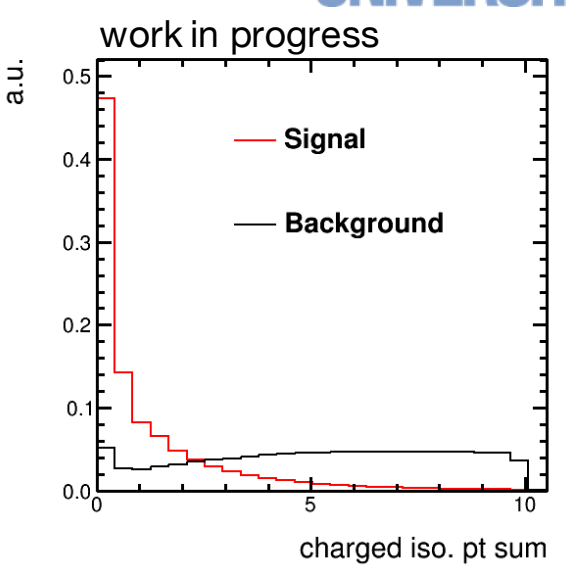
RWTH AACHEN UNIVERSITY

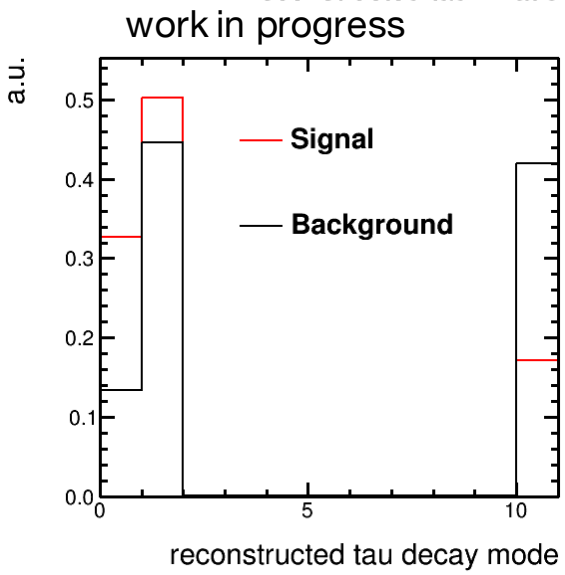
- Isolation sums:
  - → $I_{\tau,ch}$, $I_{\tau,neut}$
- Distributions and multiplicities of particles inside/outside the cone:
  - → photons P_T outside signal cone, number of photons in tau etc.
- Lifetime-related:
  - → SV, flight length, flight length significance, IP, etc.
- Tau quantities:
  - → $P_{T,\tau}$, $|\eta_{reco,\tau}|$, DM, Gottfried-Jackson-angle, energy ratio etc.

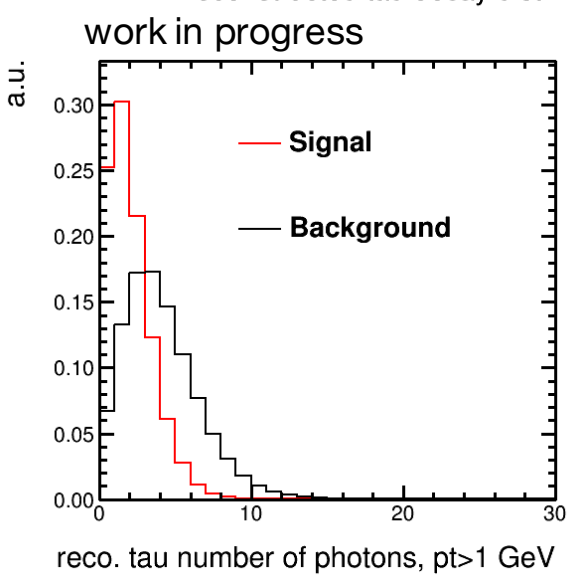
Discriminating variables

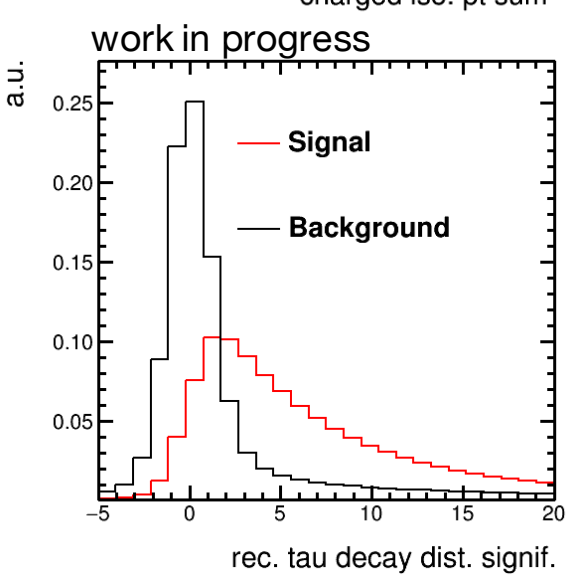




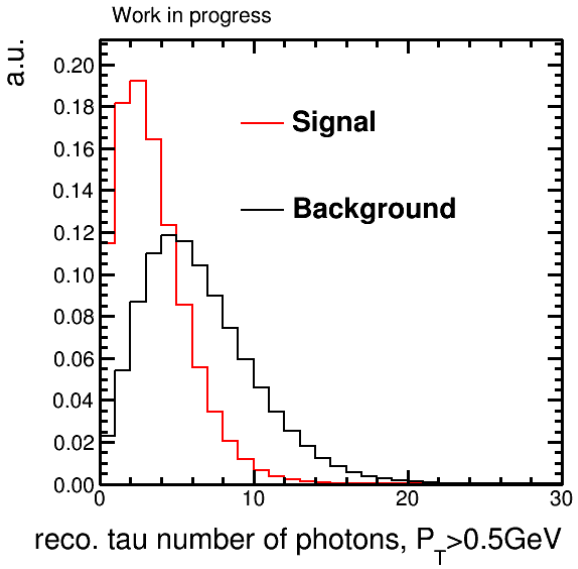


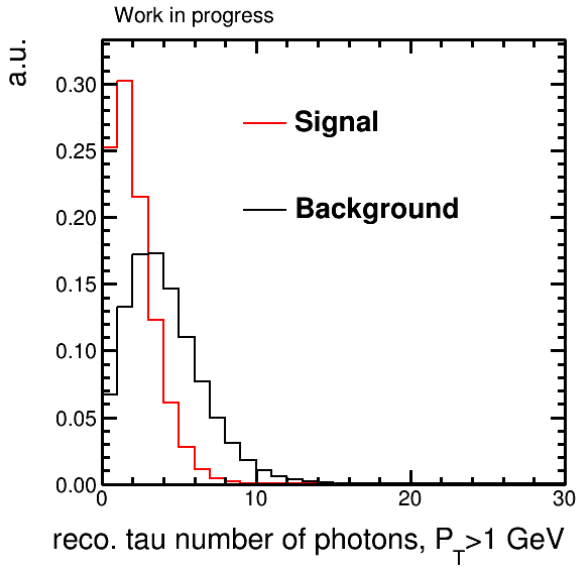


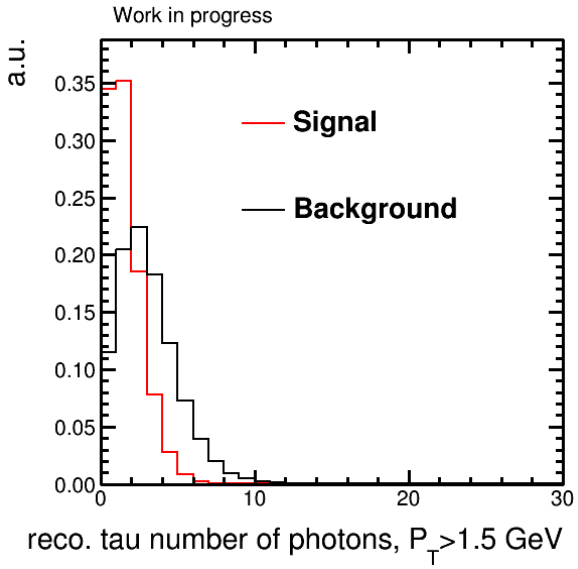




Tau ID performance in 2016







- Study of the performance on 2016 data showed that some of the tau observables had a disagreement in data/mc
  - → lower efficiency in data compared to MC
- Increasing the p_T^{γ} cut seemed an option to decrease this effect

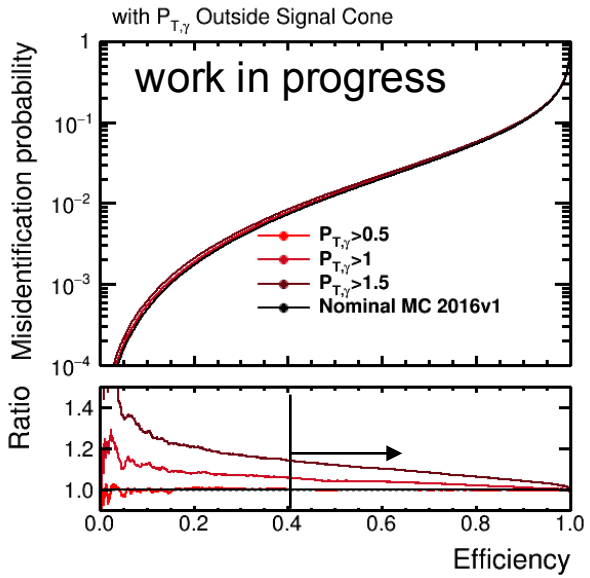
Study with an increased p_T^γ cut

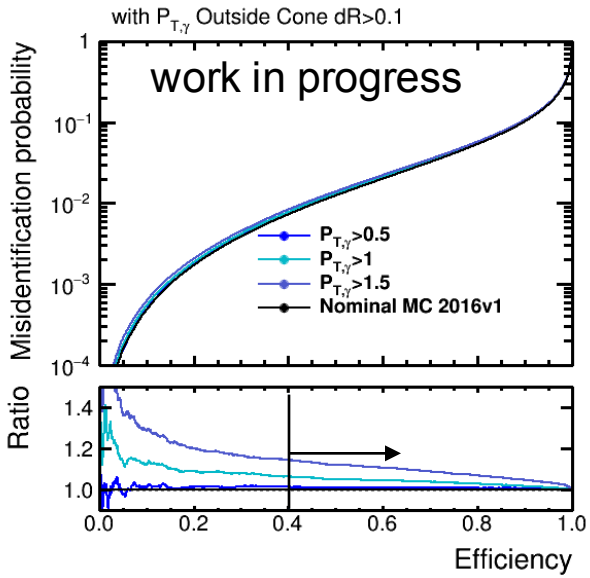
A set of trainings were performed in order to find:

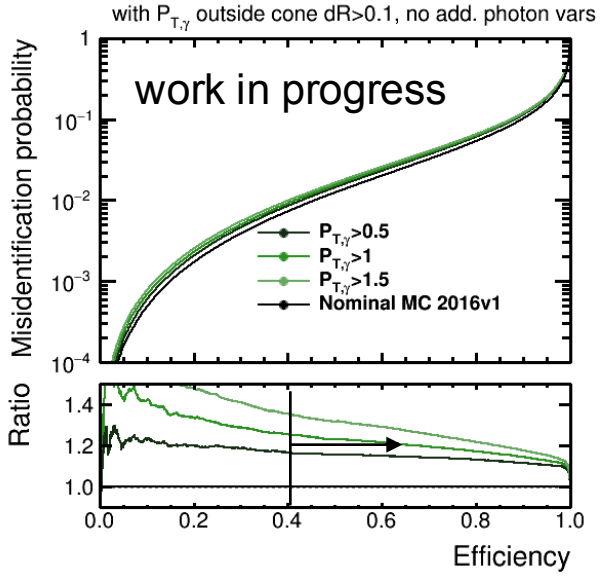
- a p_T^γ cut point
- a collection of photon-variables

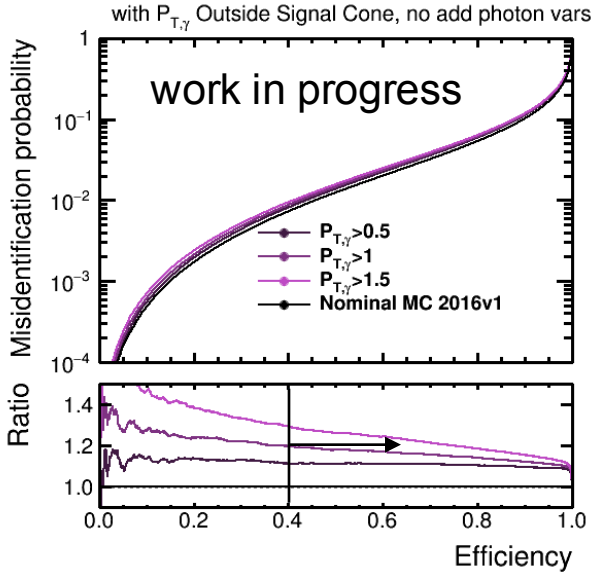
that will not decrease the performance with respect to the previous trainings

- A cut of size 1 GeV seems to be the most reasonable choice





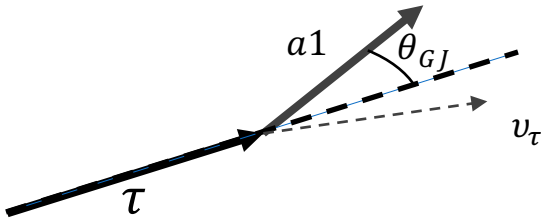


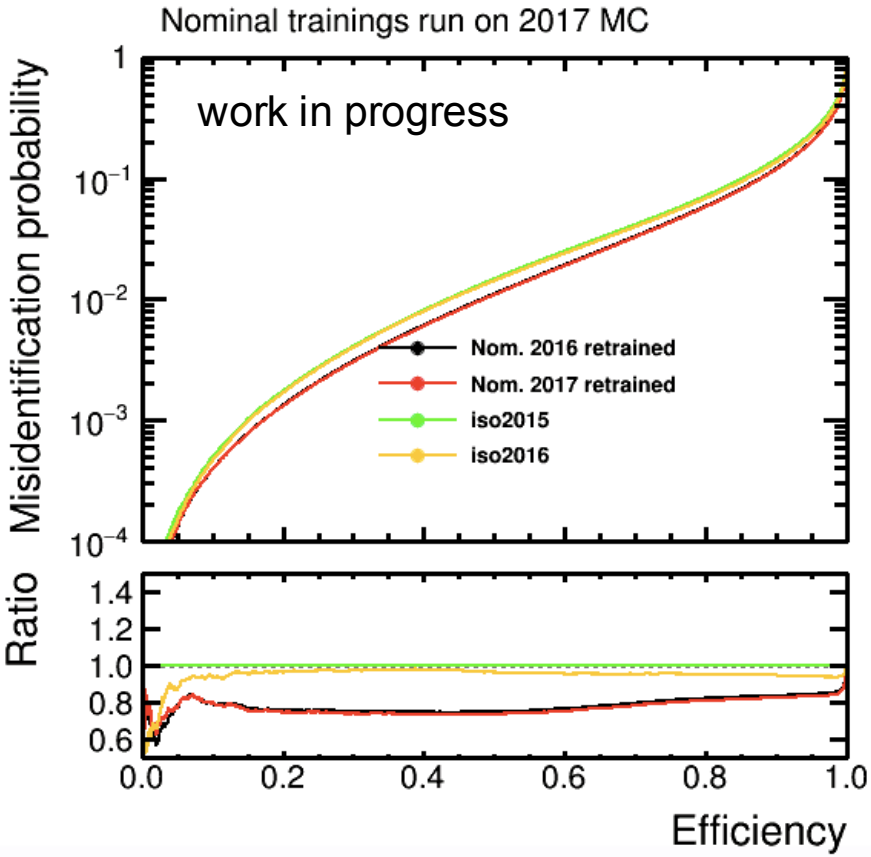


Closer look to 2017 MC trainings

- p_T^γ cut is set to 1 GeV and propagated to all photon-related observables
- Including Gottfried-Jackson-angle
  - → an angle between the flight length vector and the tau momentum

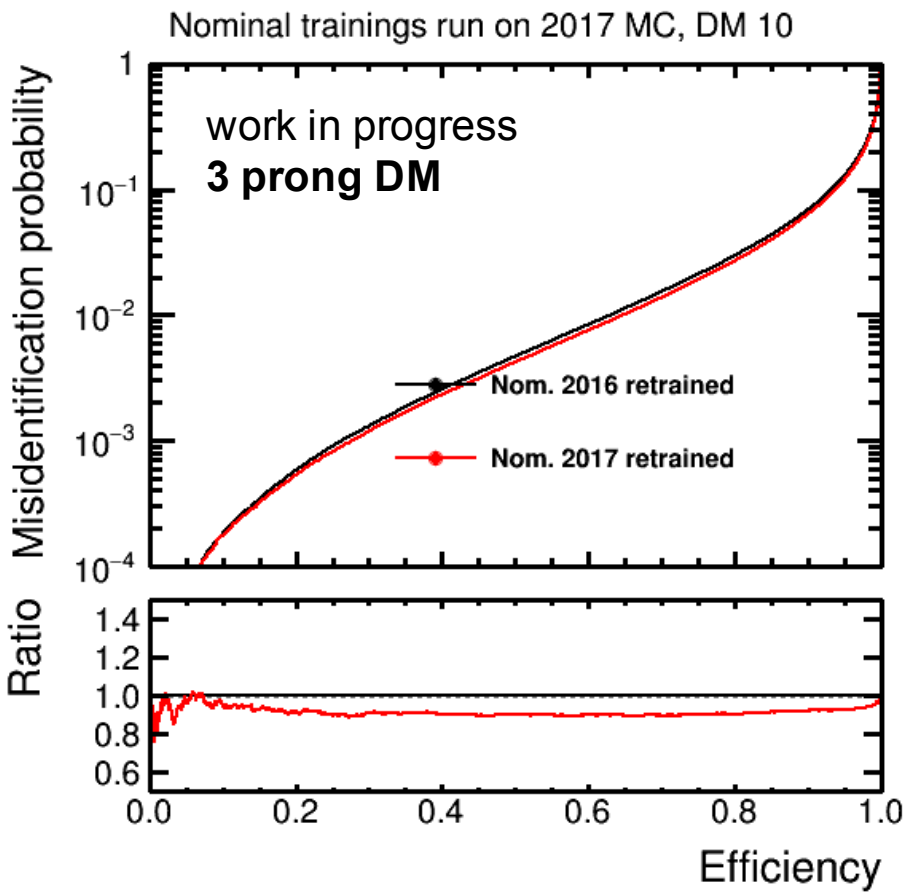
$a1$-meson resonance

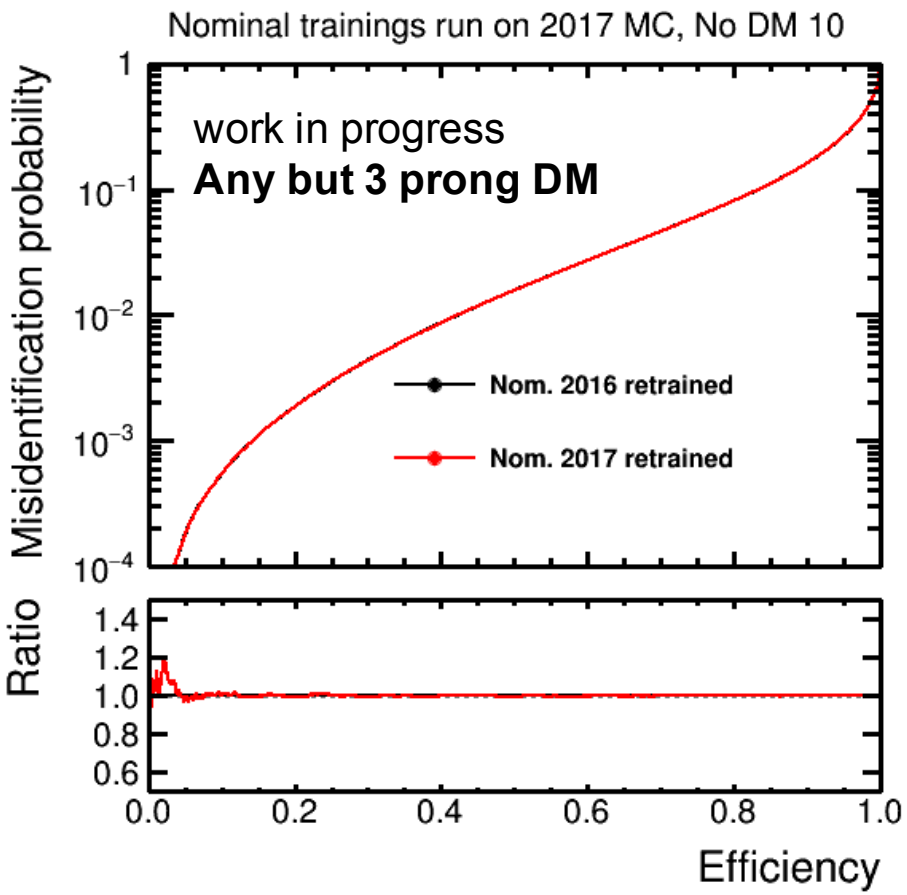




- Discarded low-rated variables

Closer look to 2017 MC trainings

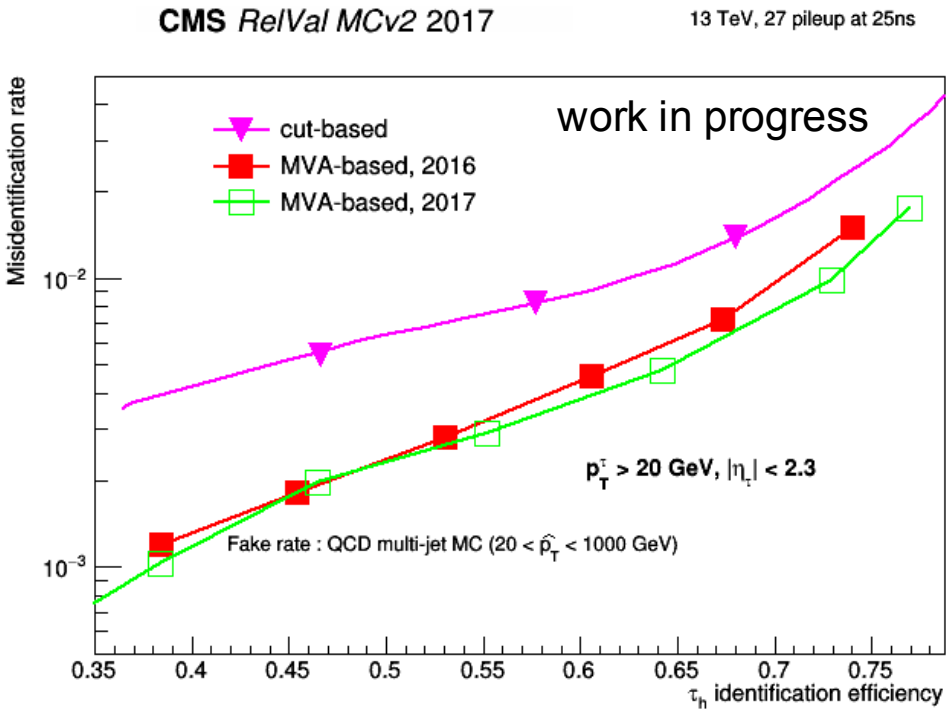




Effect of the inclusion of the GJ-angle

Conclusion

- The p_T cut on photons has been changed for the tau reconstruction to get better agreement with the data. The effect on the tau lepton reconstruction efficiency is less than 5%
- The new training was performed based on the 2017 MC which includes the upgraded pixel detector
- Performance of 2017 training shows better performance with 2017 MC
- The influence of the increasing pileup and its modelling are subject to further studies



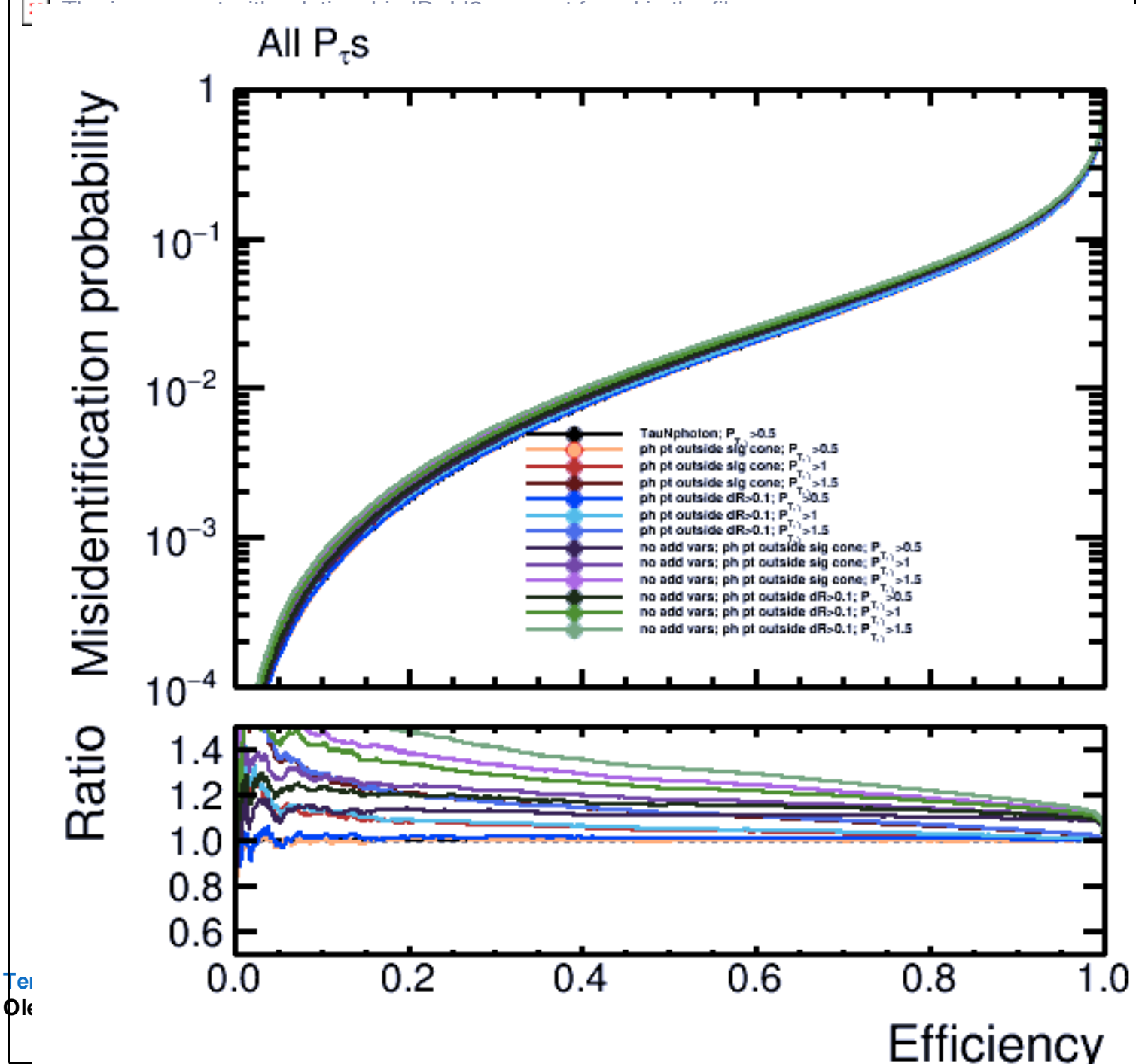
RWTH AACHEN UNIVERSITY

Appendix

Provided trainings

RWTH AACHEN UNIVERSITY

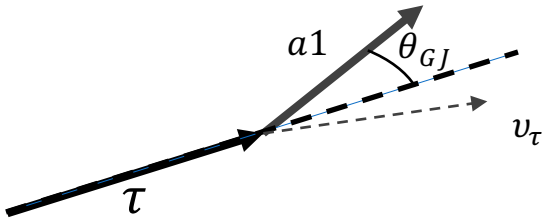
	2016 MC	2017 MC (v1)	Purpose
dR = 0.3, old DM	No ttH samples	No ttH samples (will be provided during MCv2 campaign)	For Htt analysis, meant for high hadr. activity, trained using only ttH and ttbar samples
dR = 0.5, old DM	Provided during production	Will be included in the future samples	Standard
dR = 0.3, new DM	Not requested	Not requested	-
dR = 0.5, new DM	Done & uploaded	Will be provided by the end of November	For boosted tau ID studies


All P_τs
Misidentification probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
TauNphoton; $P_{T_{\tau_1}}$>0.5
ph pt outside sig cone; $P_{T_{\tau_1}}$>0.5
ph pt outside sig cone; $P_{T_{\tau_1}}$>1
ph pt outside sig cone; $P_{T_{\tau_1}}$>1.5
ph pt outside dR>0.1; $P_{T_{\tau_1}}$>0.5
ph pt outside dR>0.1; $P_{T_{\tau_1}}$>1
ph pt outside dR>0.1; $P_{T_{\tau_1}}$>1.5
no add vars; ph pt outside sig cone; $P_{T_{\tau_1}}$>0.5
no add vars; ph pt outside sig cone; $P_{T_{\tau_1}}$>1
no add vars; ph pt outside sig cone; $P_{T_{\tau_1}}$>1.5
no add vars; ph pt outside dR>0.1; $P_{T_{\tau_1}}$>0.5
no add vars; ph pt outside dR>0.1; $P_{T_{\tau_1}}$>1
no add vars; ph pt outside dR>0.1; $P_{T_{\tau_1}}$>1.5
Ratio
1.4
1.2
1.0
0.8
0.6
0.0
0.2
0.4
0.6
0.8
1.0
Efficiency

Input variables

- $P_{T,\tau}$, $|\eta_{reco,\tau}|$, $I_{\tau,ch}$, $I_{\tau,neut}$, $P_{T,\tau,PU\ corr}$
- Photons P_T outside signal cone
- Tau decay mode, SV, flight length, flight length significance
- Number of photons in tau
- Reweighting weights for tau P_T over strip $d\eta/d\varphi/dR_{sig.}/dR_{iso.}$
- E_{ecal}/E_{hcal}
- $IP_{XY(Z)}$, sign($IP_{XY(Z)}$), $IP_{XY(Z)}$ significance
- Including Gottfried-Jackson-angle
  → an angle between the flight length vector and the tau momentum

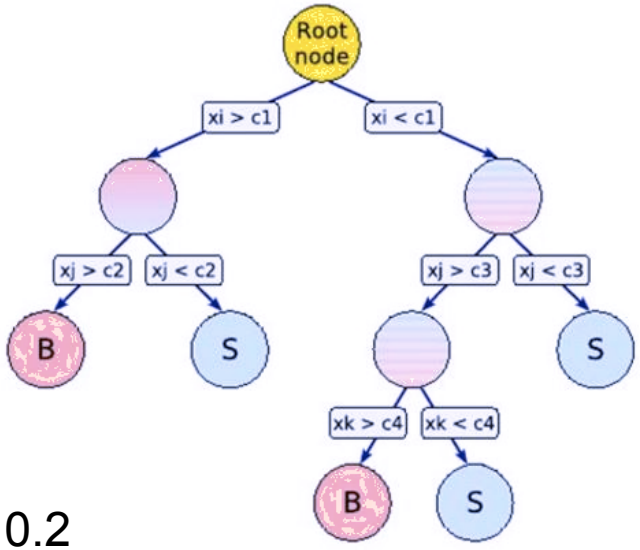
$a1$-meson resonance



MVA-based isolation

RWTH AACHEN UNIVERSITY

- Tool: TMVA
- Gradient Boost BDT
  - → Classification & regression
  - → 1000 trees in the forest
  - → Boosting type for the trees in the forest: Gradient
  - → Learning rate for GradBoost algorithm: 0.2
  - → Use only a random subsample of all events for growing the trees in each iteration
  - → Fraction of events to be used in each iteration: 0.5
  - → Separation criterion for node splitting: GiniIndex
  - → Number of grid points in variable range used in finding optimal cut in node splitting: 500
  - → Maximum depth of cell tree: 5



Samples preselecting & reweighting

- Preselecting requirements:
  - → $N_{PV} \geq 1$
  - → $|z_{PV} - z_{vtx,\tau}| < 0.4$ [mm]
  - → tau's leading PF charged hadron $P_T > 1$ GeV
  - → $\sum_{charged\,cand.}^{iso.cone} P_T < 10$ GeV
  - → Reconstructed DM
- Input samples are reweighted (P_T/η) in order to not to discriminate based on the phase space