Timing studies for JetMET

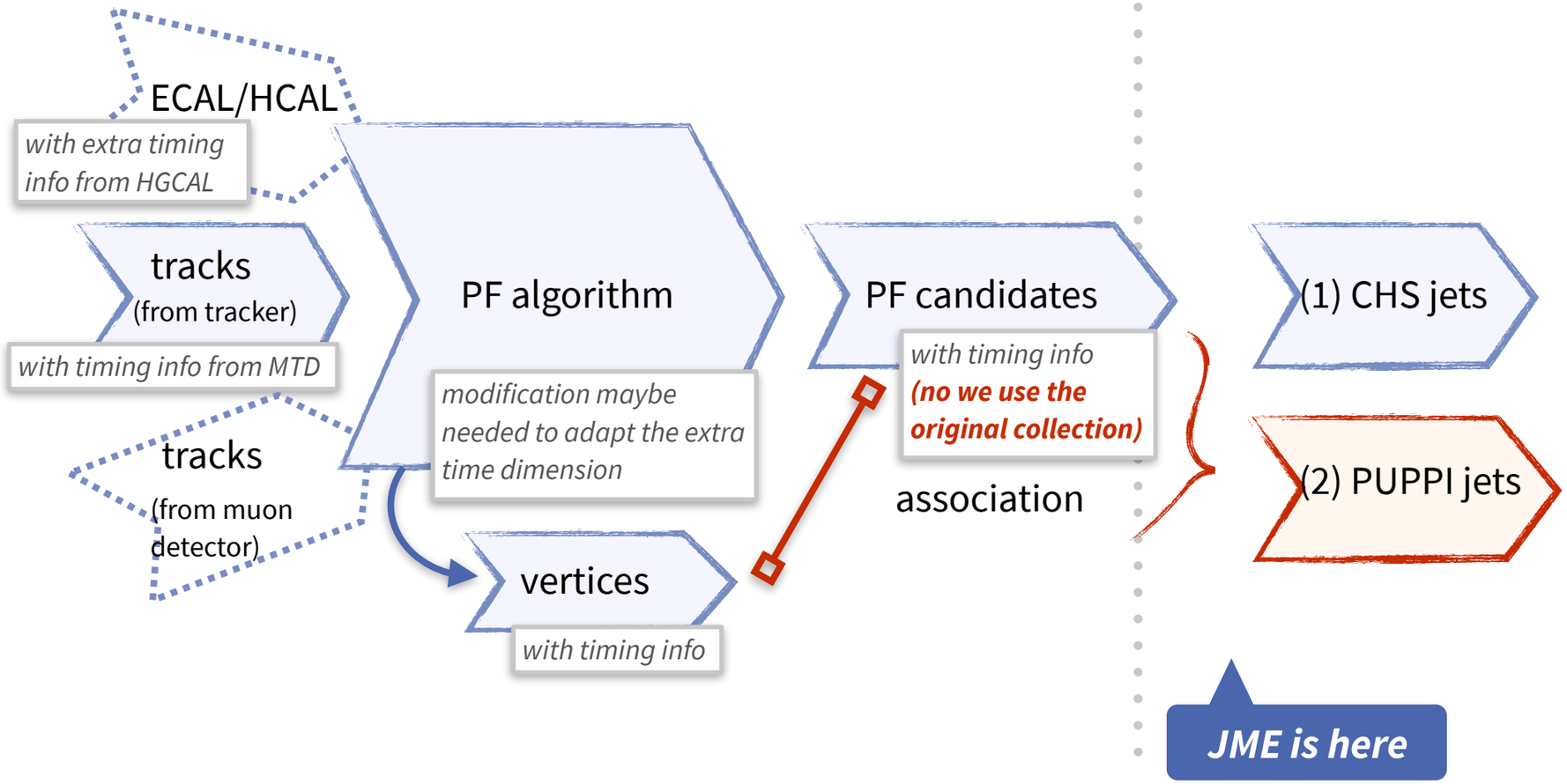
Congqiao Li *for JME group*

UPSG Future LLP and Timing: First Workshop

22 September, 2021

Introduction

- ➔ Timing information will be available during Phase-II upgrade
  - ❖ MTD allows 4D tracking and vertexing (3D + timing info); HGCAL also provides timing measurements
- ➔ Jet and MET algorithms need updates based on better handling of PF candidates
  - ❖ update the PF reconstruction strategy using the additional time dimension of tracks → update JME algorithms (should be optimal in performance)
  - ❖ modify the JME algorithm by using existing PF candidates with 4D vertices (preliminary results in this talk)



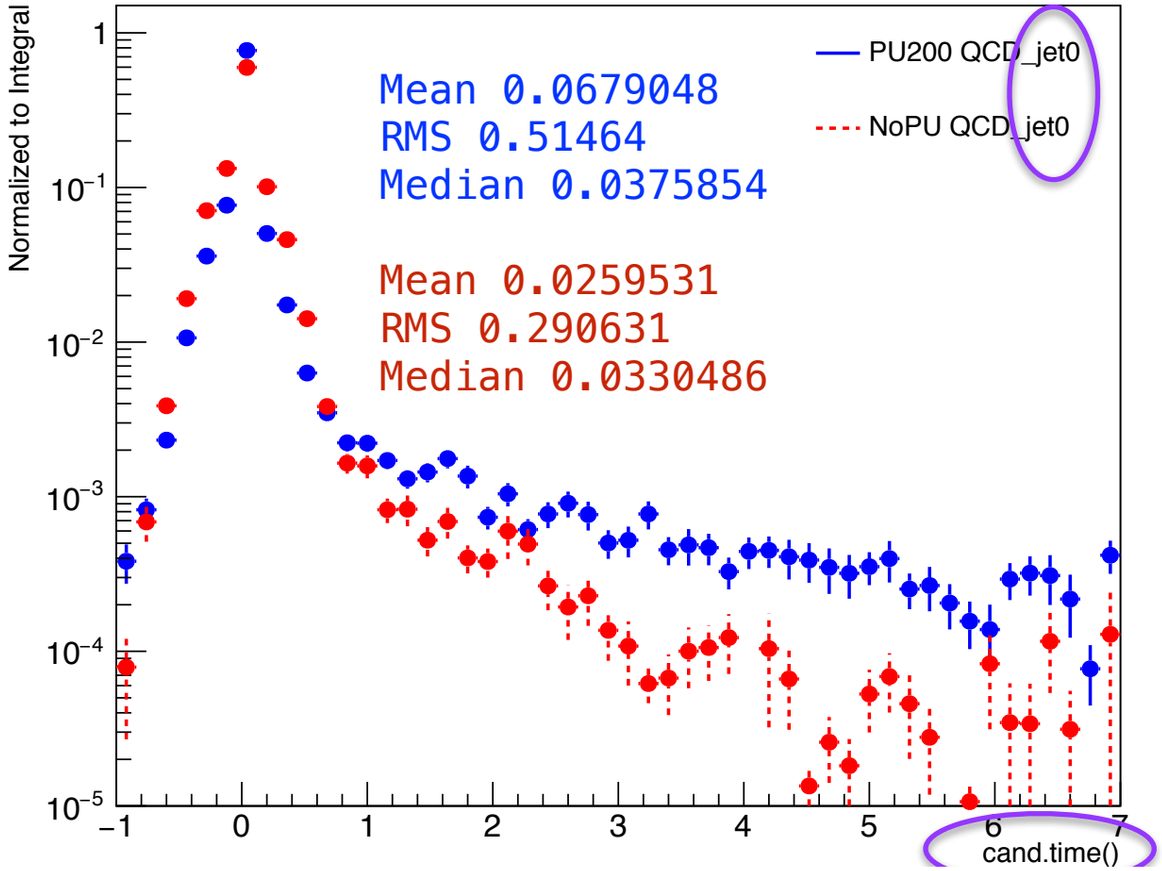
- ➔ In this talk, we introduce:
  - ❖ timing performance with PU jet ID [details in slides from Debarati Roy]
  - ❖ PUPPI with 4D vertices [details in slides from Anna Benecke]

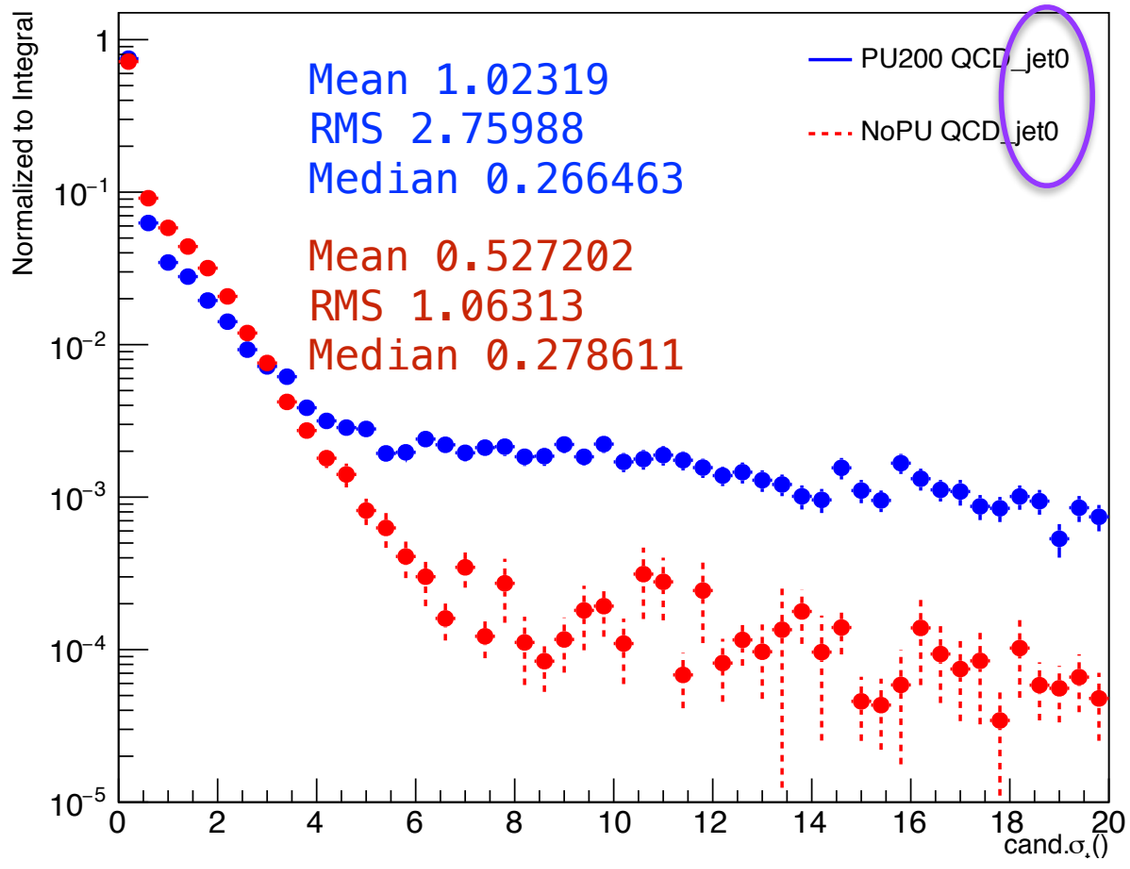
Timing for PU samples

See Debarati Roy's slides for details

➔ Timing performance on candidates with PU200 / NoPU samples

- samples: `/RelValQCD_Pt15To7000_Flat_14/CMSSW_11_1_0_pre6-PU25ns_110X_mcRun4_realistic_v3_2026D49PU200-v1/MINIAODSIM`
  `/RelValQCD_Pt15To7000_Flat_14/CMSSW_11_1_0_pre6-110X_mcRun4_realistic_v3_2026D49noPU-v1/MINIAODSIM`
- timing variables
  - t_{cand}
  - $\sigma_{t,\text{cand}} = |t_{\text{cand}} - t_{\text{PV}}| / \sqrt{err^2_{t,\text{cand}} + err^2_{t,\text{PV}}}$

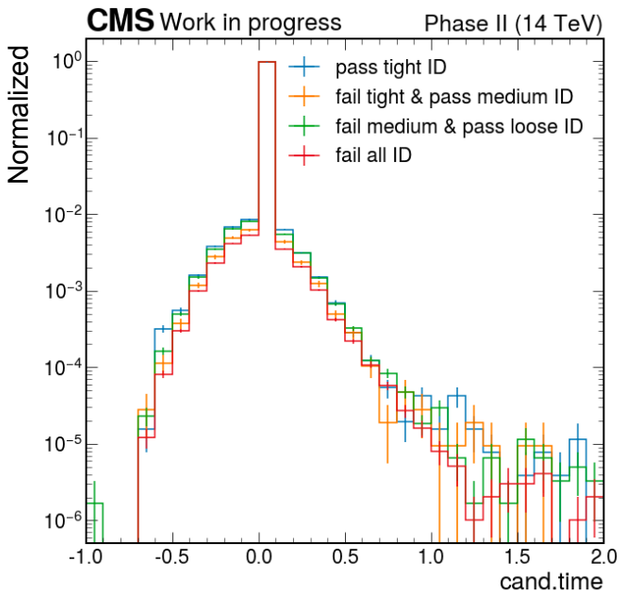


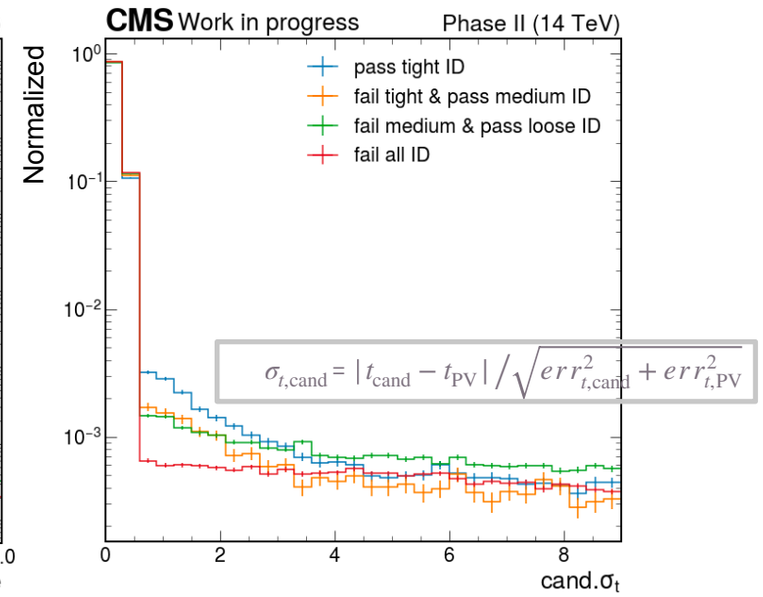


Timing for PU jet ID

➔ Check the timing information for PU jet IDs

- ❖ use PU200 sample in p.5
- ❖ collect all candidates in CHS jets with pT<50 GeV
- ❖ categorize based on PU jet IDs (tight/medium/loose/fail all)
  - ▸ note: pass tight PU ID → likely to be real jets



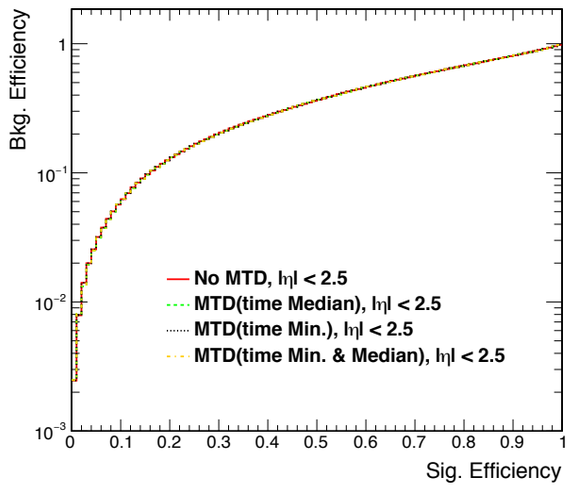


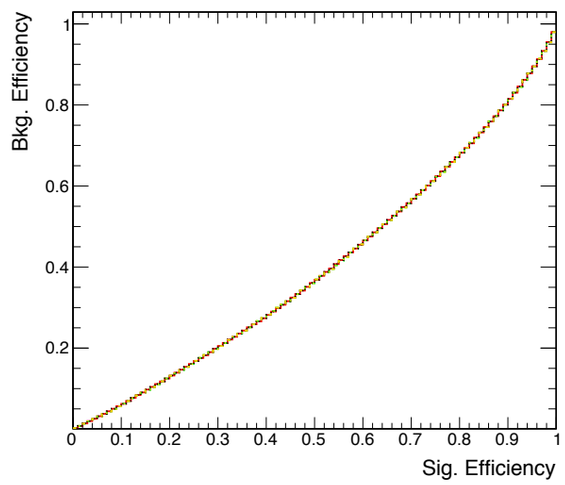
➔ Candidate time and σ_t does not follow what we expect

- ❖ although similar with Debarati's finding [slides], with a different PU categorization scheme

➔ Retrain the BDT for PU jet ID in CHS jets

See Debarati Roy's slides (MTD DPG) for details





BDT training (signal => Real jets, background => PU jets, NTrees=800, BoostType=Grad)

➔ Results from Debarati also show that a retraining of the BDT (giving the PU jet ID score) does not show improvement

➔ Reason may be the ineffectiveness and high fake rate of the current 4D vertex collection

PUPPI for 4D vertices — setup

See Anna Benecke's slides for details

➔ Samples

Global tag: 111X_mcRun4_realistic_T15_v1
Era: eras.Phase2C9
Geometrie: Configuration.Geometry.GeometryExtended2026D49_cff

/QCD_Pt_300to470_TuneCP5_14TeV_pythia8/Phase2HLTTDRSummer20ReRECOMiniAOD-NoPU_111X_mcRun4_realistic_T15_v1-v1/GEN-SIM-DIGI-RAW-MINIAOD
/QCD_Pt_300to470_TuneCP5_14TeV_pythia8/Phase2HLTTDRSummer20ReRECOMiniAOD-PU140_111X_mcRun4_realistic_T15_v1-v1/GEN-SIM-DIGI-RAW-MINIAOD
/QCD_Pt_300to470_TuneCP5_14TeV_pythia8/Phase2HLTTDRSummer20ReRECOMiniAOD-PU200_111X_mcRun4_realistic_T15_v1-v1/GEN-SIM-DIGI-RAW-MINIAOD

➔ Match PF candidates with 4D/3D vertices

- 4D vertices from "offlineSlimmedPrimaryVertices4D"
- test on the additional option (for 4D) to use timing info in the track-vertex association criteria

$$dist. = \frac{d_{z,\mathrm{cand}} - d_{z,\mathrm{vert}}}{\sqrt{err^2_{d_{z,\mathrm{cand}}} + err^2_{d_{z,\mathrm{vert}}}}} + \frac{t_{\mathrm{cand}} - t_{\mathrm{vert}}}{Reso_{t_{\mathrm{cand}}}}$$

➔ Adopt a simple PUPPI tune

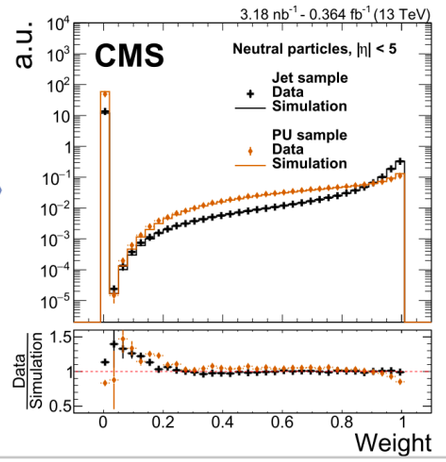
- "dz scheme": assign tracks to the closest vertex in *dist.*, plus requiring dz < 0.1 && dzsig < 5 tracks to the assigned vertex (for 4D collection, further require dt < 0.1)

recap on PUPPI

tracks (charged particles) ***associated with vertices*** → *calculate the α variable for all neutral particles* → *calculate signed χ^2 based on α; assign weights to neutral particles*

(based on distributions of charged tracks associated with LV)

(charged particles have weight either 0 (pileup) or 1 (associated to LV))

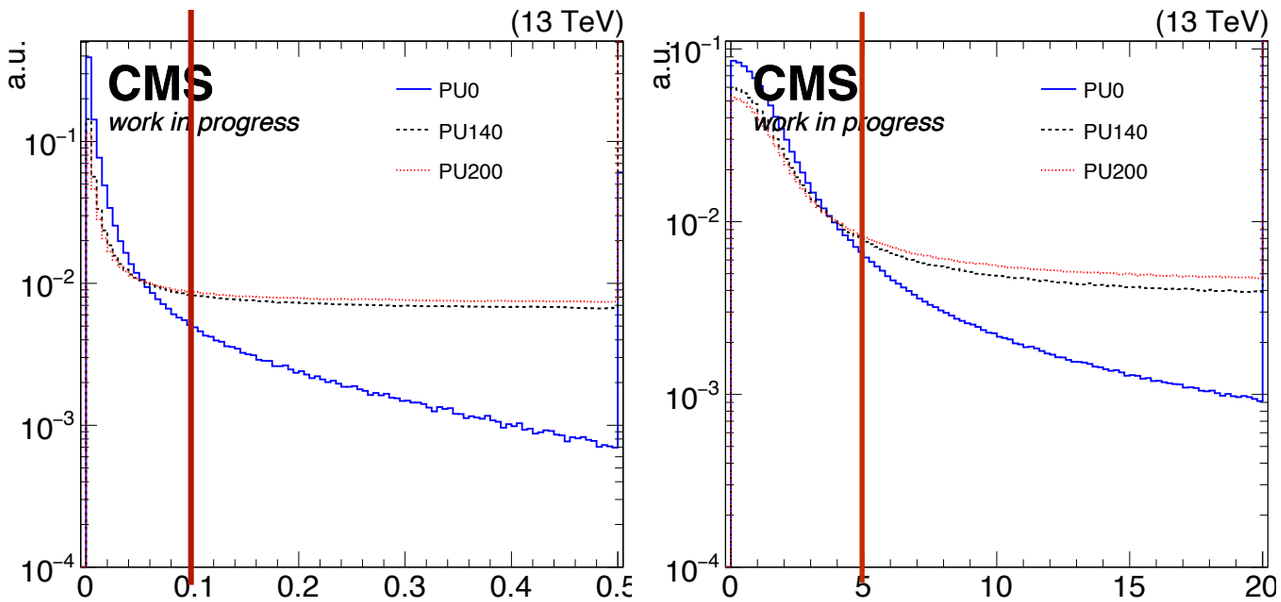


(scale candidate 4-momentum by the weight)

Effect of PUPPI "dz scheme"

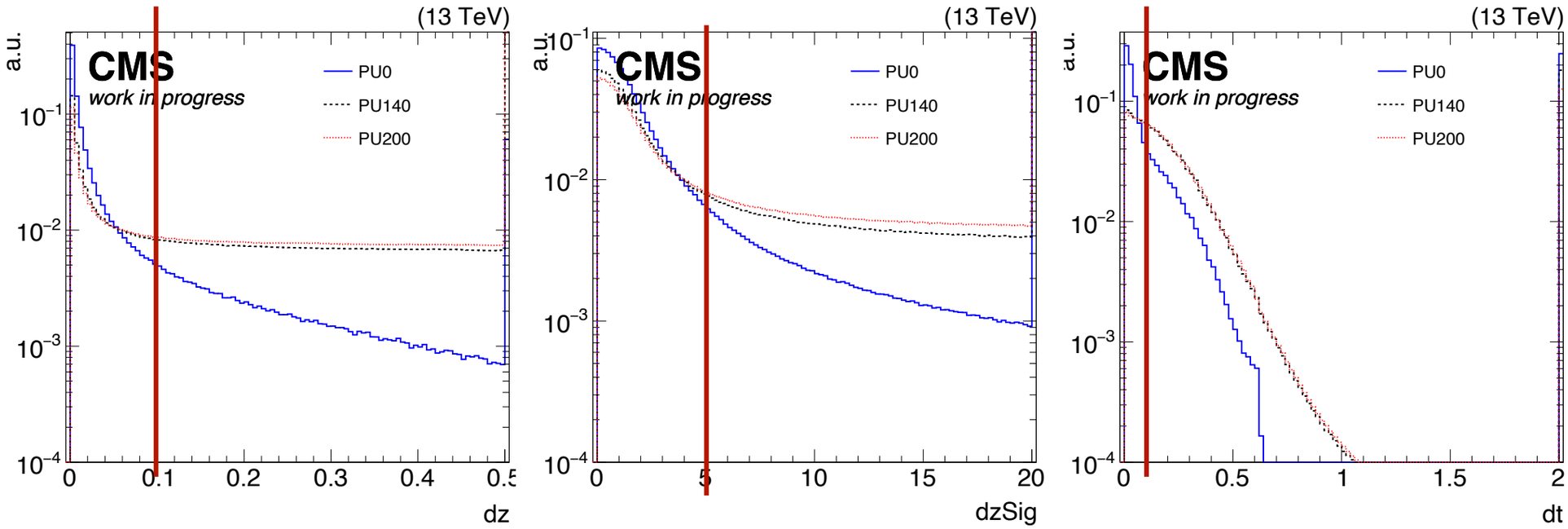
See Anna Benecke's slides for details

with 3D vertex collection



- ➔ Variables are shown without PUPPI weights
- ➔ dz and dzSig effectively filters the pileup candidates (which are wrongly assigned to LV)

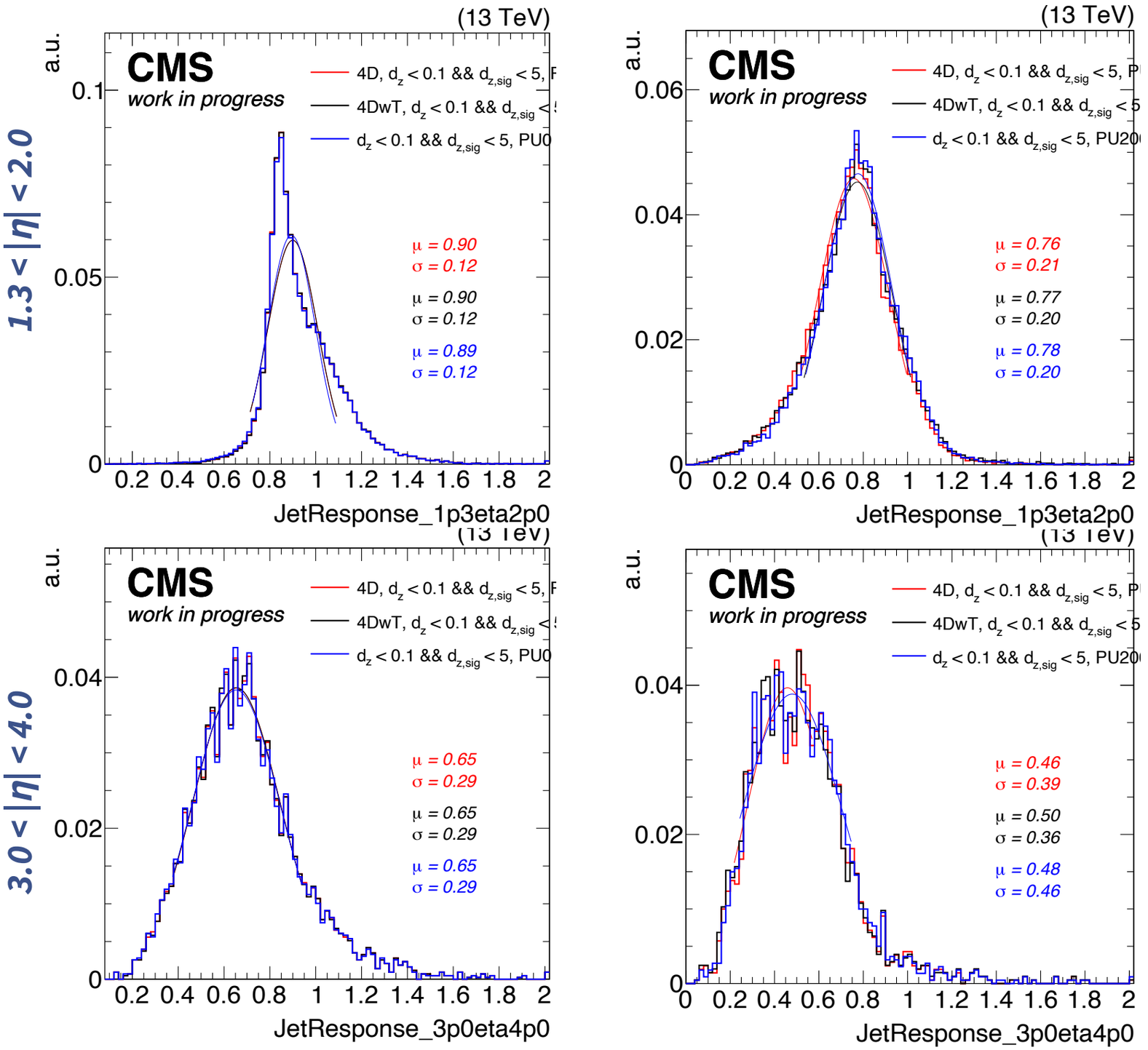
with 4D vertex collection



PUPPI for 4D vertices

See Anna Benecke's slides for details

noPU | **PU200**



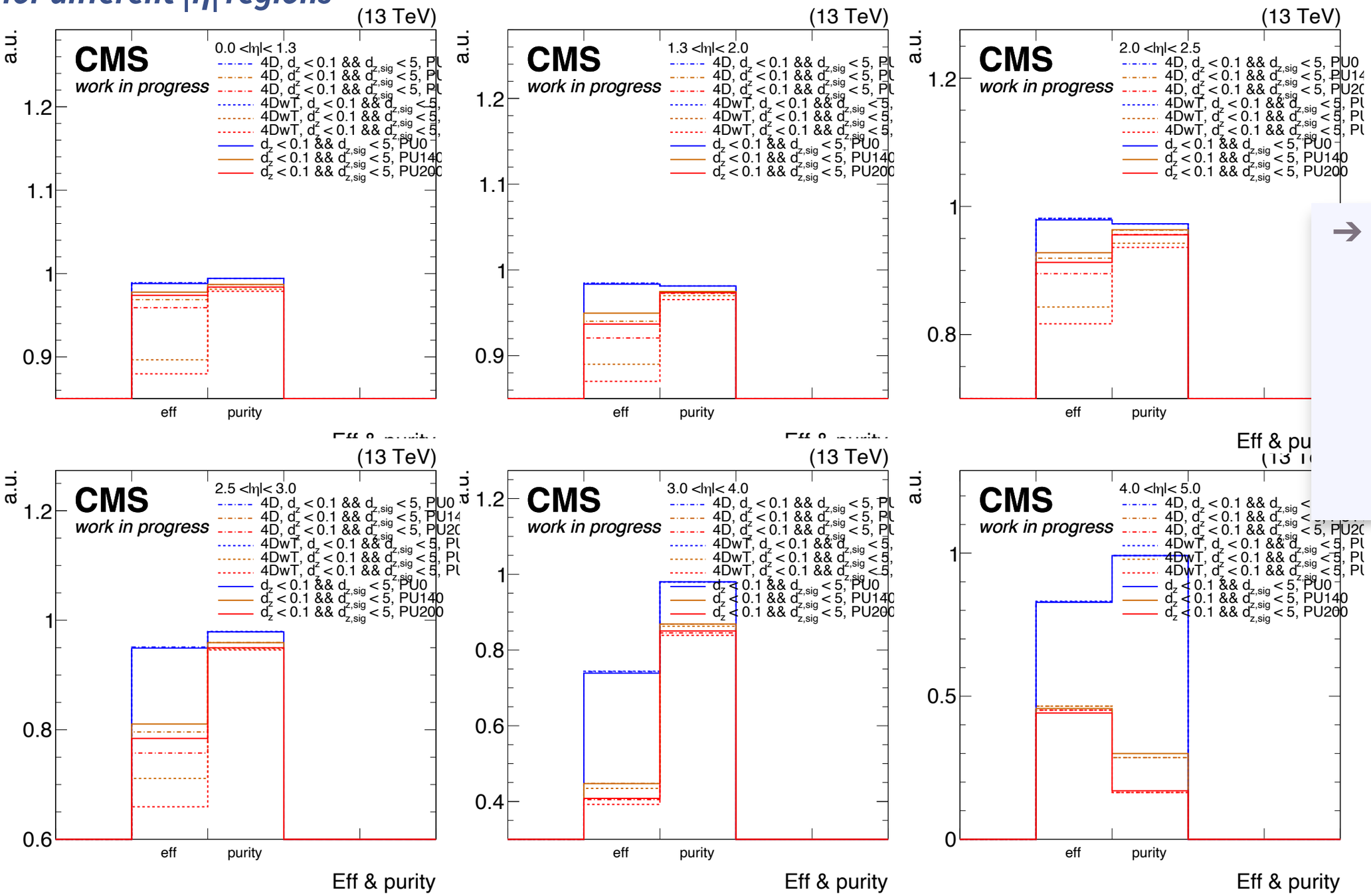
jet response: jet energy / matched gen-jet energy

➔ No significant change from
- ❖ 3D → 4D vertex collection
- ❖ with/without adding the time info in the association criteria

Efficiency/purity of PUPPI jets

See Anna Benecke's slides for details

for different |η| regions

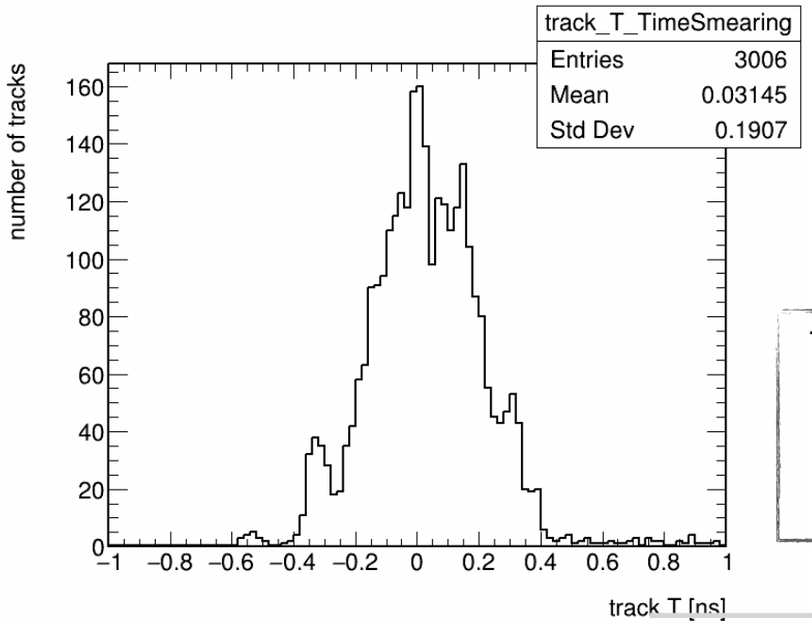


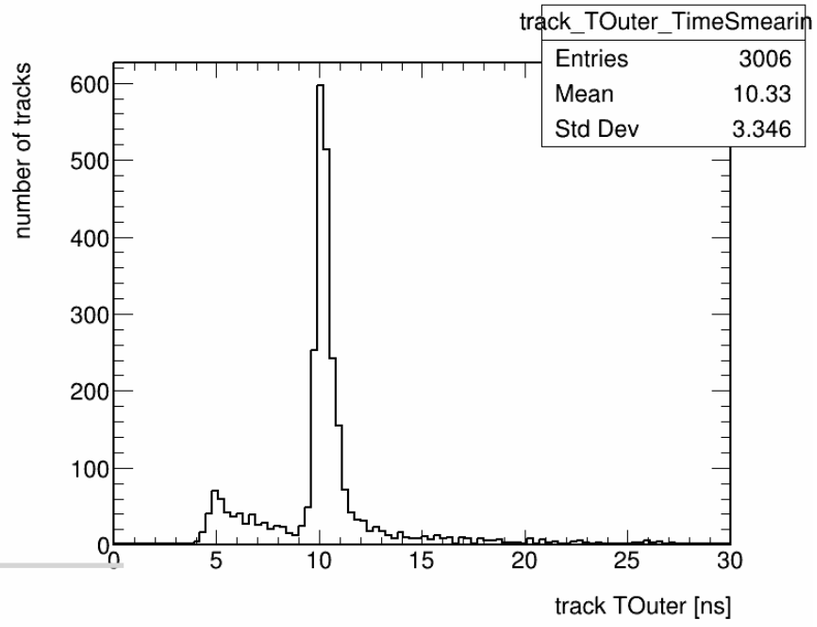
➔ 4D vertex collection has a lower efficiency compared to 3D, especially when applying a timing requirement

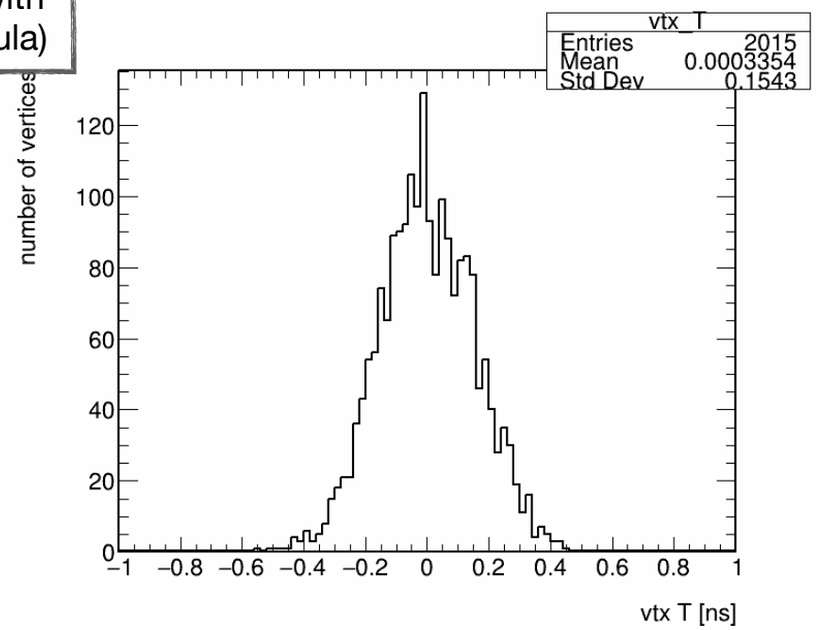
efficiency: *fraction of gen-jets (pT>30) that have a matched reco PUPPI jet (pT>20)*
purity: *fraction of reco PUPPI jets (pT>30) that have a matched gen-jet (pT>20)*

PUPPI studies with Delphes

Work in progress by Anna Albrecht







- Signal: VBF_HH_bbbb
- 200 PU (delphes card CMS_PhaseII_200PU_v03.tcl)
  - ZVertexSpread = 0.25 m
  - TVertexSpread 0.8 ns
- TrackSmearing : trackResolutionCMS.tcl
- TimeSmearing
  - 30 ps resolution
  - Eta dependent resolution possible in new module (not implemented yet)

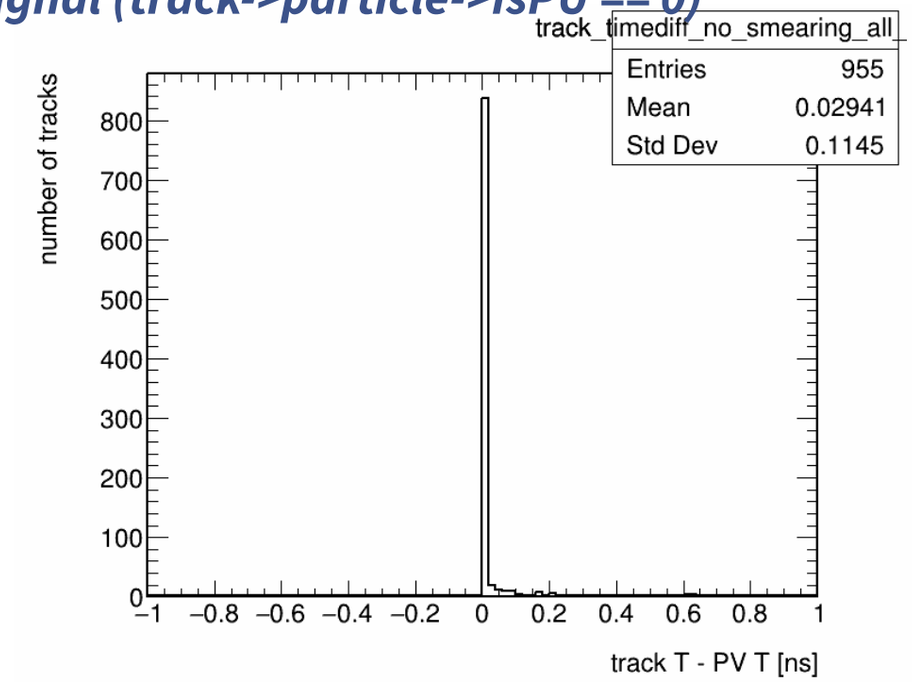
GenJet Collections:
GenJets (clustered from signal genparticles)
GenJetsPU (clustered from PU genparticles)

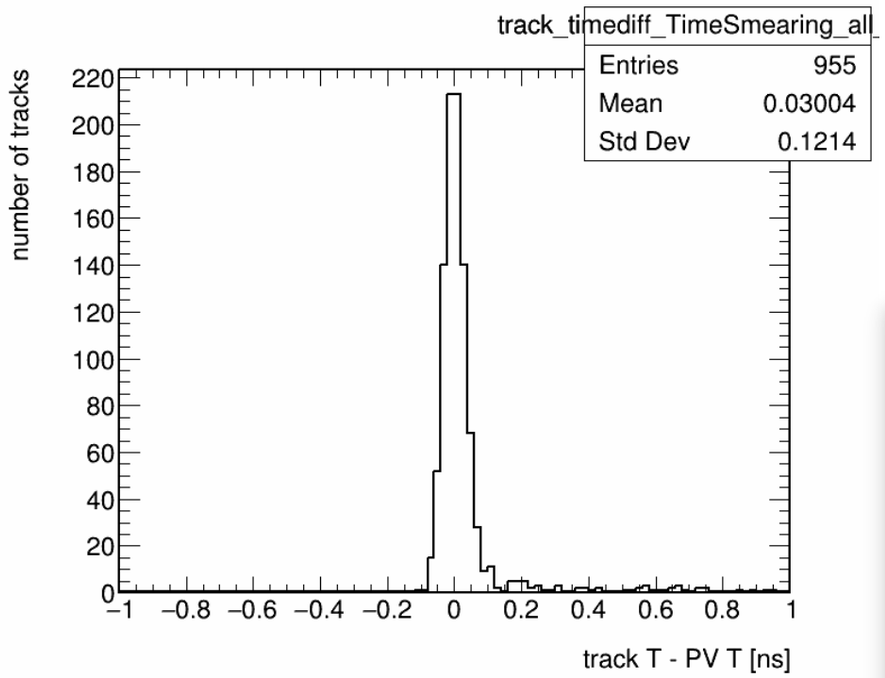
Jet Collections:
CHS jets, PUPPI jets

PUPPI studies with Delphes (II)

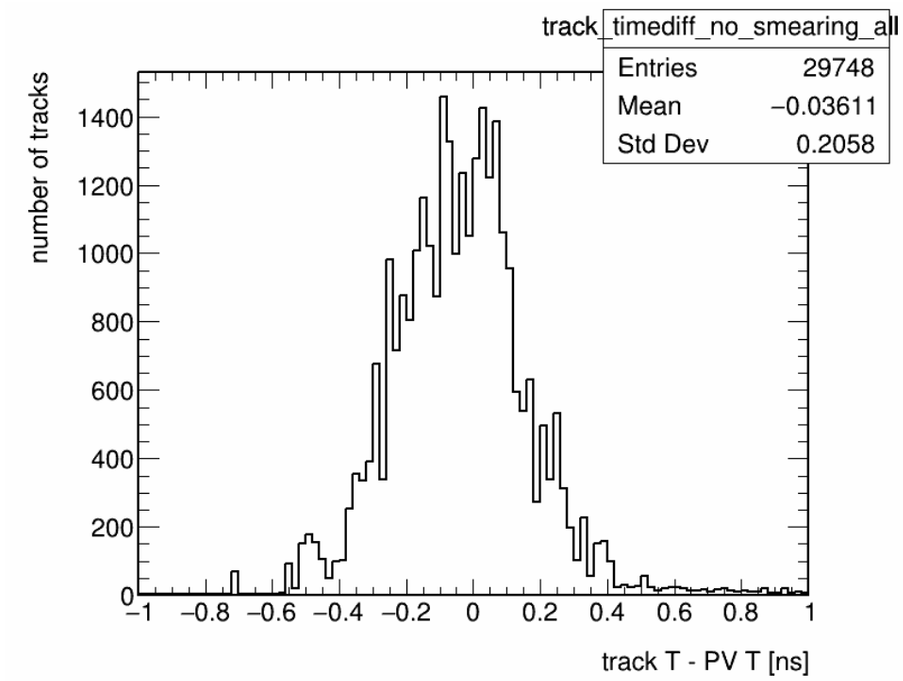
Work in progress by Anna Albrecht

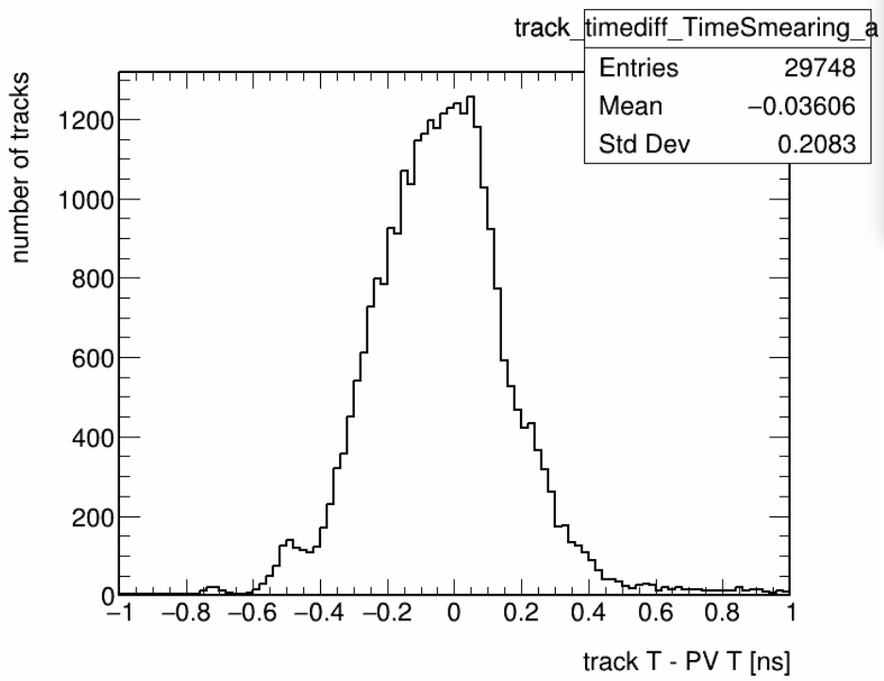
Signal (track->particle->IsPU == 0)





PU (track->particle->IsPU == 1)





➔ Using Delphes to distinguish the real tracks vs. fake (pileup)

- ❖ PU tracks show wider $t_{track} - t_{PV}$ distribution as expected

➔ Further variable studies on Delphes instruct future PUPPI tuning

Summary

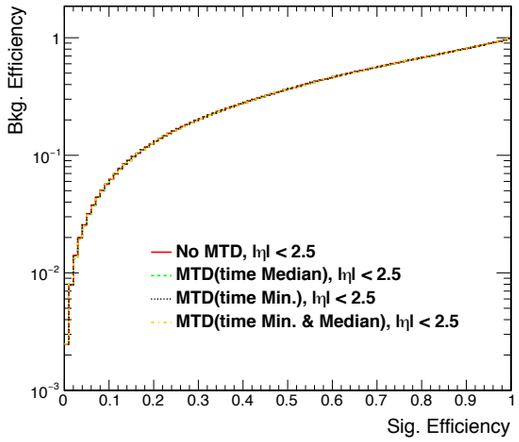
- ➔ Additional timing info contradicts the trend of real/PU jets differences
  - ❖ need more study, although future pileup mitigation will go on top of PUPPI
- ➔ Preliminary attempt to use PUPPI based on 4D track-vertex association
  - ❖ adopt the “dz scheme” for the track-vertex association for ease of analysis
  - ❖ do not see improvement in 4D association, for using and without using the time as the matching criteria
  - ❖ use Delphes analysis as a guide in understanding timing info and for further tuning of PUPPI
- ➔ Further studies:
  - ❖ tune the PUPPI and better handle the time in track-vertex association
  - ❖ consider the timing info in the recent ML-related pileup mitigation effort

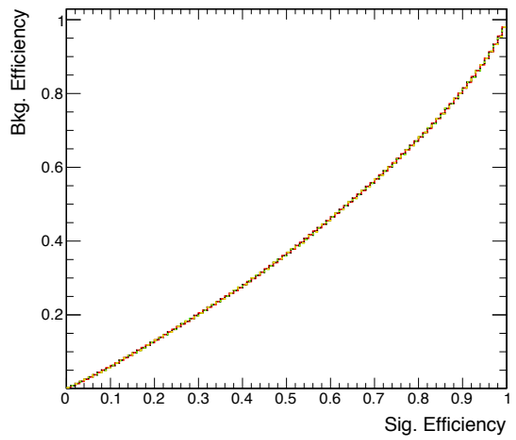
Backup

Timing for PU jet ID BDT retraining

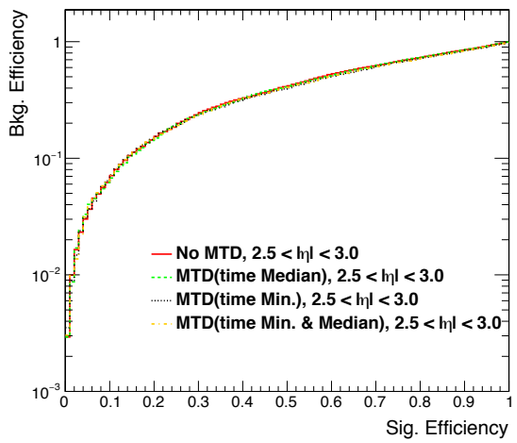
See Debarati Roy's slides (MTD DPG) for details

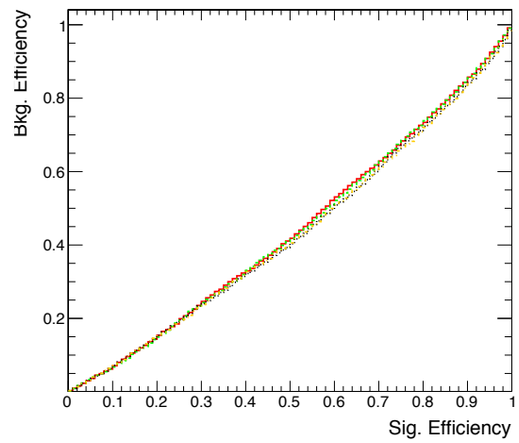
➔ Retrain the BDT for PU jet ID in CHS jets

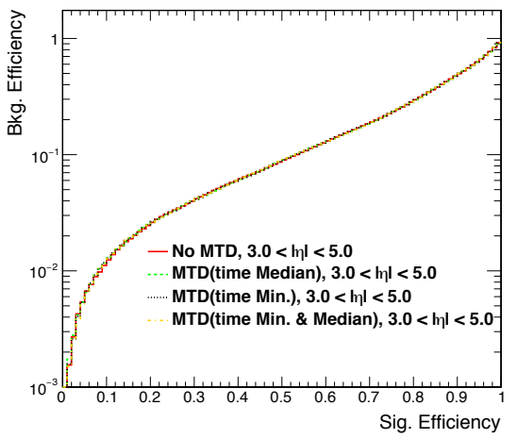


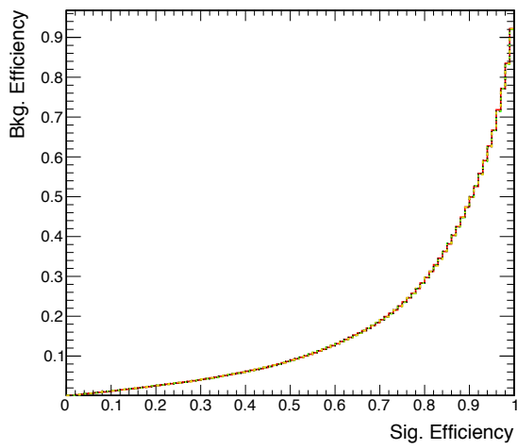


BDT training (signal => Real jets, background => PU jets, NTrees=800, BoostType=Grad)









BDT input for CHS jets PU jet ID. Jet time info included as additional variable

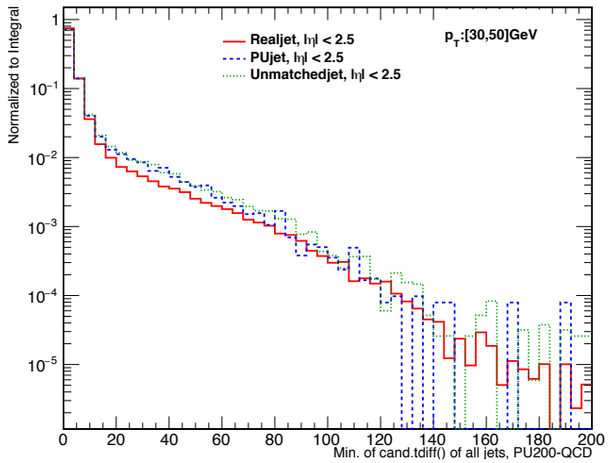
β /beta	fraction of transverse momentum of charged particles associated to the primary vertex, defined as $\frac{\sum_{i \in V} p_{Ti}}{\sum_i p_{Ti}}$ where i iterates over all the PF particles in the jet
$n_{vertices}$	number of vertices in the event
$\langle \Delta R^2 \rangle$ /dR2Mean	p_T^2 average weighted by square distance of jet constituents from the jet axis : $\frac{\sum_i \Delta R^2 p_{Ti}^2}{\sum_i p_{Ti}^2}$
frac01, frac02, frac03 etc=> f_{ringX}, X = 1, 2, 3, and 4	fraction of p_T of the constituents ($\sum p_{Ti}/p_T^{jet}$) in the region $R_i < \Delta R < R_{i+1}$ around the jet axis, where $R_i = 0, 0.1, 0.2$, and 0.3 for X=1, 2, 3, and 4
p_T^{lead}/p_T^{jet} /jetR	transverse momentum fraction carried by the leading PF candidate
$p_T^{l.\,ch.}/p_T^{jet}$ /jetRchg	transverse momentum fraction carried by the leading charged PF candidate
$\lvert\vec{m}\rvert$ /pull	pull magnitude, defined as $\lvert(\sum_i p_T^i \lvert r_i \rvert \vec{r}_i)\rvert / p_T^{jet}$ where $\vec{r}_i$ is the direction of the particle i from the direction of the jet
NParticles/ N_{total}	number of PF candidates
NCharged/ $N_{charged}$	number of charged PF candidates
majW/ σ_1	major axis of the jet ellipsoid in the η-ϕ space
minW/ σ_2	minor axis of the jet ellipsoid in the η-ϕ space
p_T^D	jet fragmentation distribution, defined as $\sqrt{\sum_i p_{T_i}^2} / \sum_i p_{Ti}$

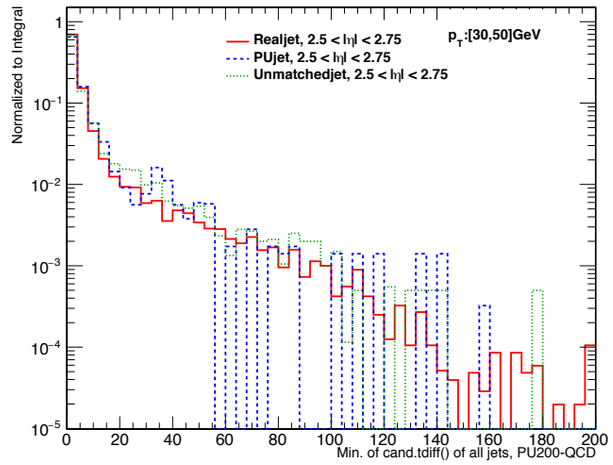
Timing for PU jet ID: different PU categorization

See Debarati Roy's slides (MTD DPG) for details

Categories as real(Real), pileup(PU) & unmatched jets with PU200 sample with stricter phase space :

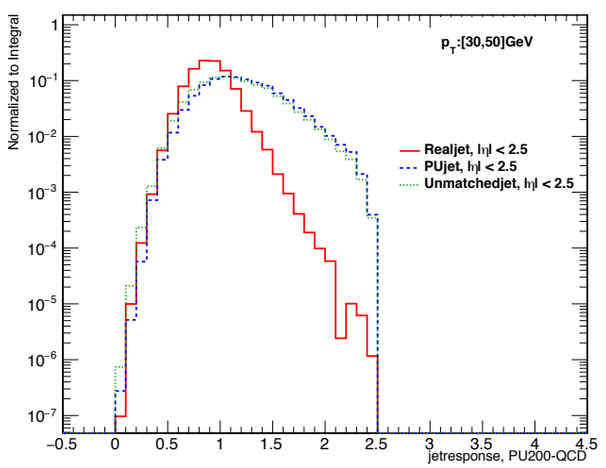
- Defined real jet as : $p_T > 30$ GeV (reco level), $p_T > 20$ GeV (gen level), $|\eta| < 5.0$, $\Delta R < 0.1$, |partonflavour| != 0.
- Defined pileup jet as : $p_T > 30$ GeV (reco level), $p_T > 20$ GeV (gen level), $|\eta| < 5.0$, $\Delta R > 0.4$, |partonflavour| == 0.
- Defined unmatched jet as : $p_T > 30$ GeV (reco level), $p_T > 20$ GeV (gen level), $|\eta| < 5.0$.
- 4 $|\eta|$ slices further considered : < 2.5, $2.5 < |\eta| < 2.75$, $2.75 < |\eta| < 3.0$, $3.0 < |\eta| < 5.0$ that further fed to BDT.
- 2 p_T slices considered : 30 to 50 GeV, 30 to 100 GeV. [New Results]

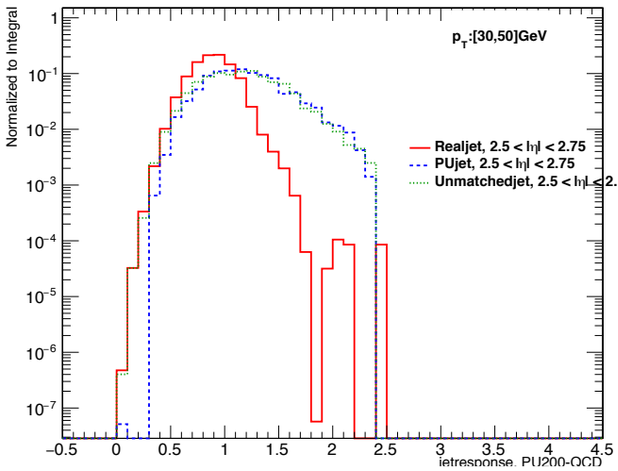




In more central region variable more effective

➔ Small difference seen in the timing variable; clear distinction show in the jet response



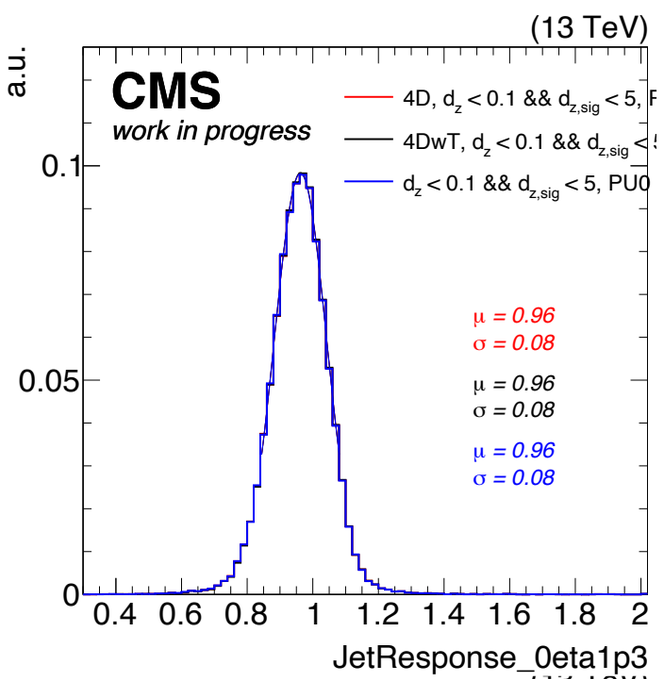


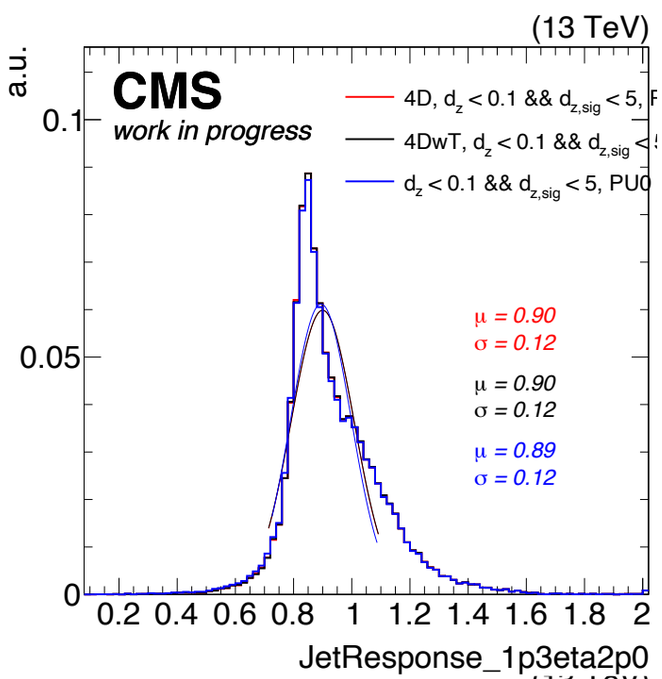
Jet response shows clear distinction

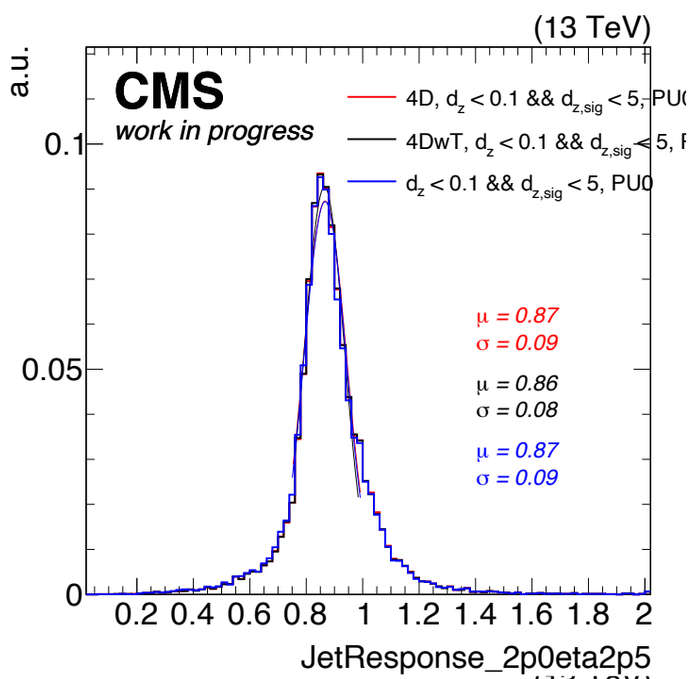
PUPPI for 4D vertices: jet response

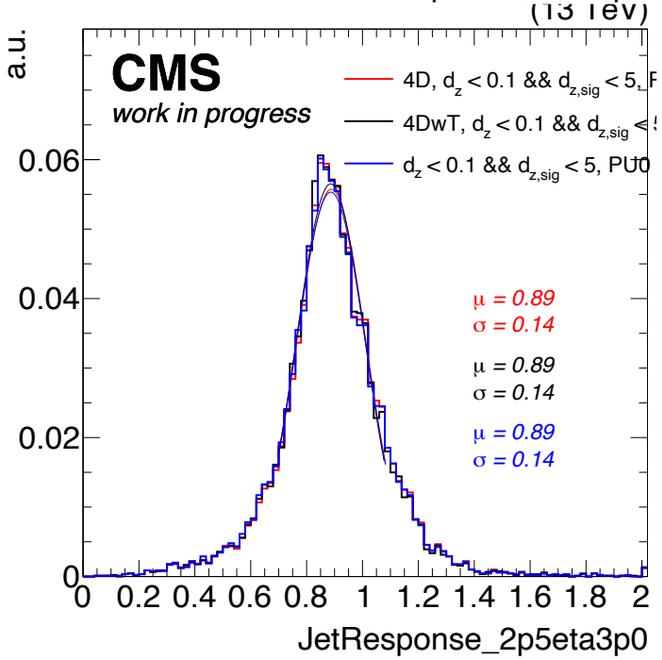
See Anna Benecke's slides for details

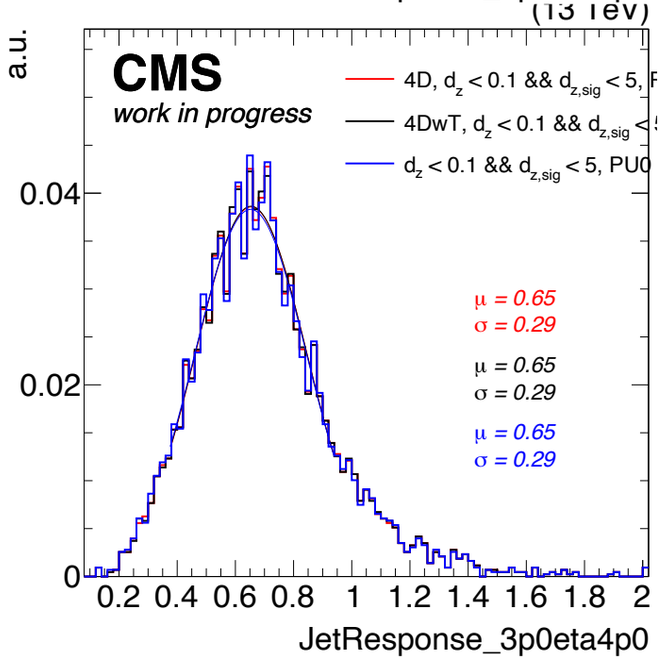
Jet response 0PU

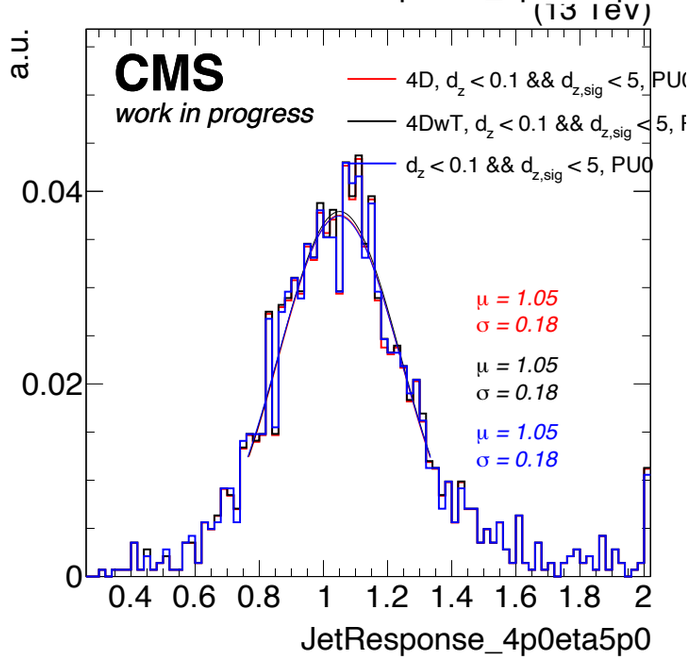












PUPPI for 4D vertices: jet response (II)

See Anna Benecke's slides for details

Jet response 200PU