Tau Seed and Hadronic Tau Substructure in ATLAS

Chenzheng Zhu, Yang Liu, Yuchen Cai

Institute of High Energy Physics, Chinese Academy of Science

The 6th China LHC Physics Workshop (CLHCP2020)

07/11/2020

Outline

- Tau seed study
  - Topological cluster jet seeds & EM Particle-flow jet seeds
  - Reconstruction efficiency
  - Estimate on p_T & η resolution, decay mode classification efficiency

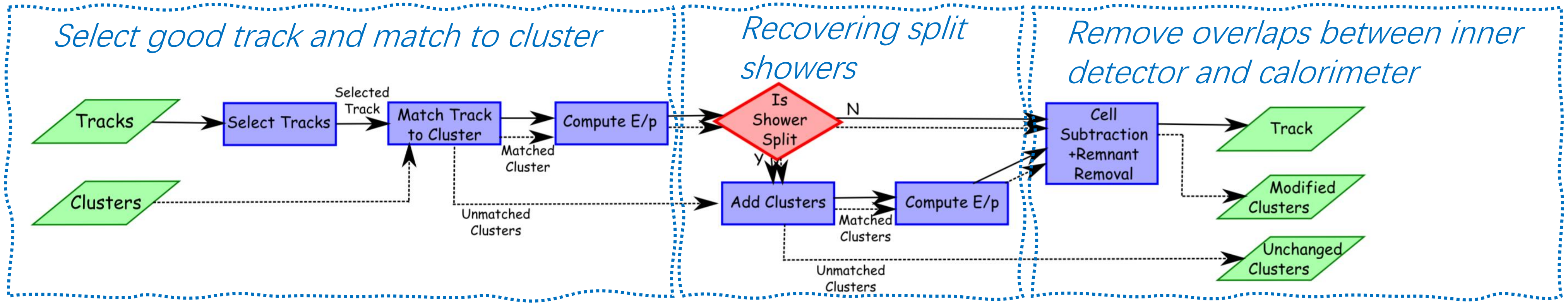
- Hadronic Tau Substructure
  - tau substructure reconstruction efficiency
  - Precision TES/TER measurement

Tau Seed Study

Introduction

- Tau Reconsturction:
  - jet → jet seeding → vertex & track → tau candidates

- jet seeds
  - Topo-clusters: Topological clusters of calorimeter cells
  - Particle-flow: calorimeters clusters + inner detector tracking information



  - CERN-EP-2017-024 : (PFlow jet) Better energy, angular resolution in the central region and Robust against pile-up

- Improve the resolution. Especially the low-pT tau performance $p_T < 20$ GeV.

$$\frac{\sigma(E)}{E} = \frac{50\,\%}{\sqrt{E}} \oplus 3.4\,\% \oplus \frac{1\,\%}{E}$$

calorimeters

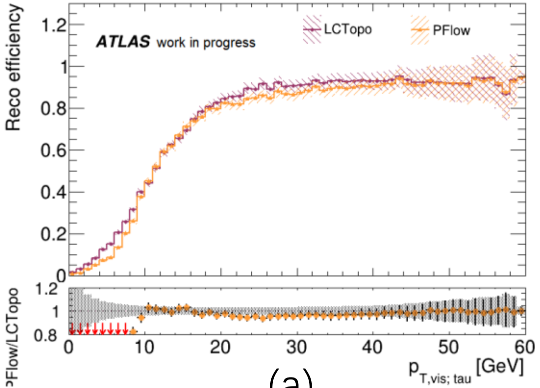
$$\sigma\left(\frac{1}{p_{\mathrm{T}}}\right) \cdot p_{\mathrm{T}} = 0.036\,\% \cdot p_{\mathrm{T}} \oplus 1.3\,\%$$

tracker

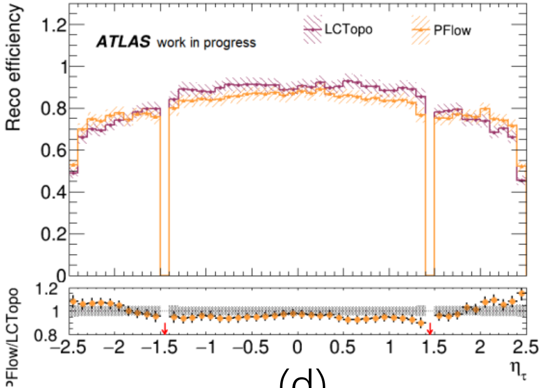
- ILC, CMS …

Results

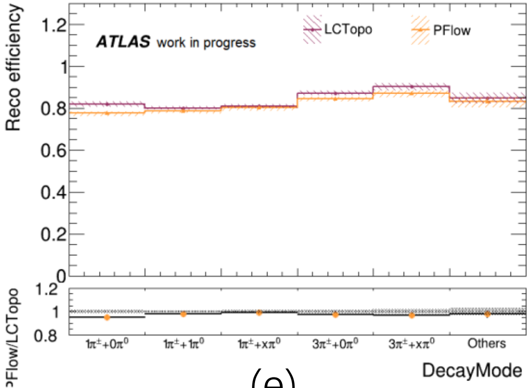
- $\epsilon_{reco} = \frac{N_{Taujets\ matched\ Truth}}{N_{Truth\ Tau}}$



(a)

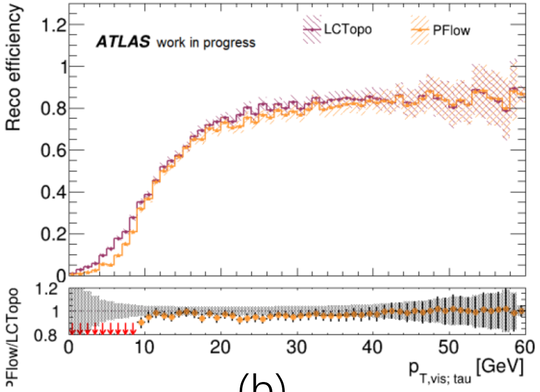


(d)

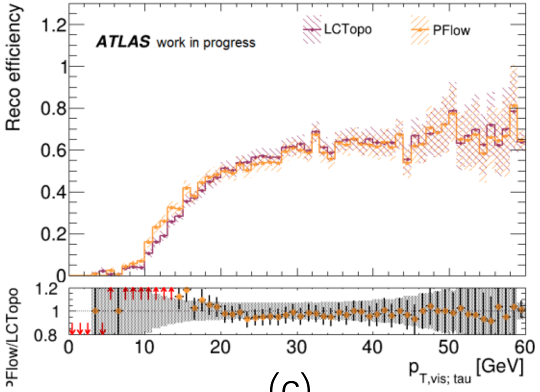


(e)

- 1 track matched to 1 prong & 3 track matched to 3 prong



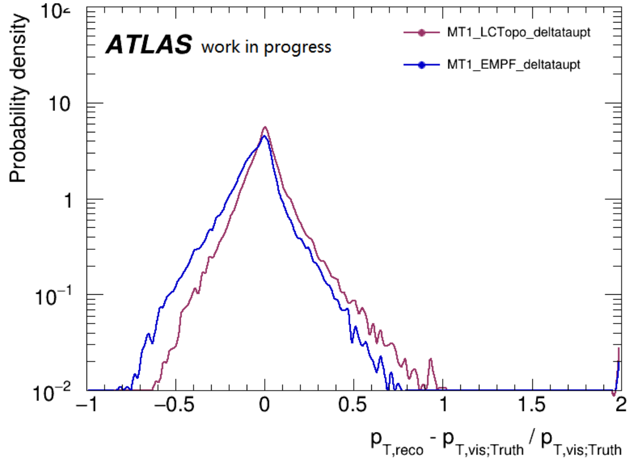
(b)

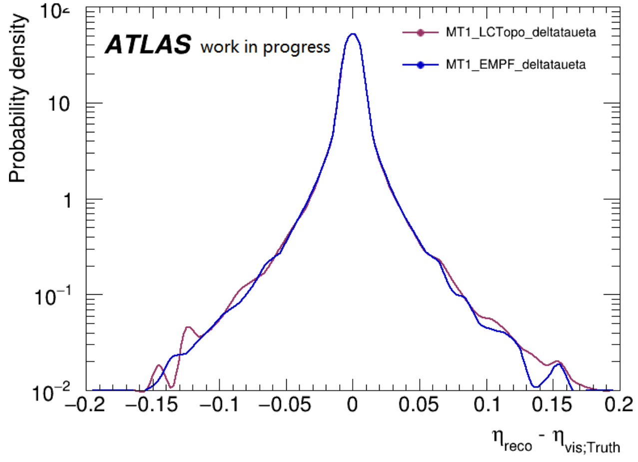


(c)

- Different performance in $p_{T\,vis} \in (10,20)\ GeV$ and $p_{T\,vis} \in (20,40)\ GeV$
- Within the expectation:
  - EMPFlow is better at low pT region
- Without the expectation:
  - EMPFlow is quiet worse at higher pT region

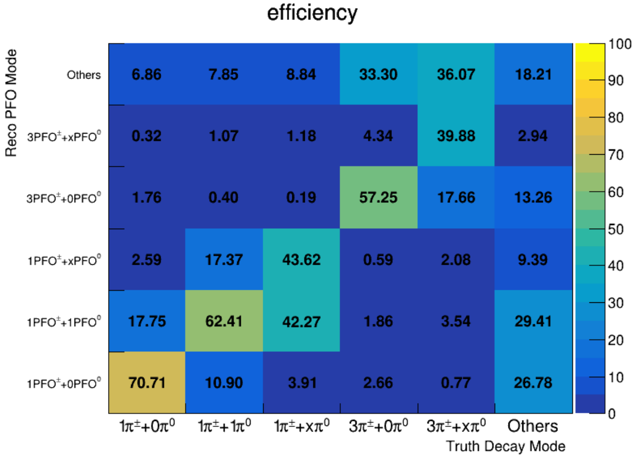
Resulation & Decay classification matrix



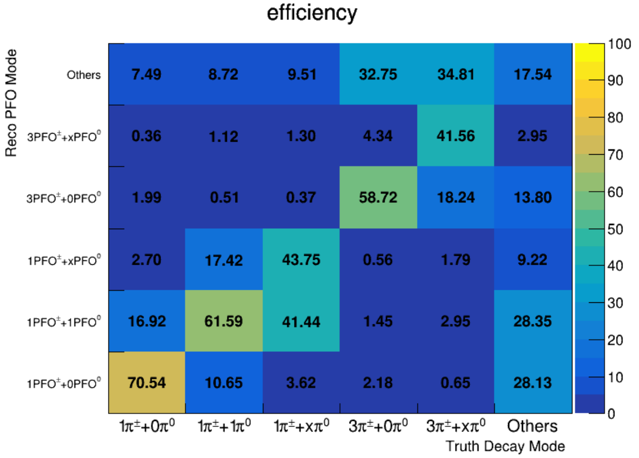


- gross estimation:
  - $\frac{p_{T_{reco}} - p_{T_{vis;Truth}}}{p_{T_{vis;Truth}}}$: $p_{T_{reco}}$ is taujet's p_T
  - $\eta_{reco} - \eta_{vis;Truth}$: η_{reco} is taujet's η

- A shift in p_T with nearly the same width.
- Nearly same in eta.



- LCTopo



- EMPFlow

- Estimation
  - Xaxis: Truth tau decay mode
  - Yaxis: Charged PFOs and π^0 PFOs from jet*
  - Not normalized because 2/4/more… charged pion
- EMPF in 3 p is slightly better

Conclusion

- By full reconstruction samples:
  - Particle Flow seeds perform well in the end-cap regions, 3-prong situations, and high neutral pion energy fraction. The inefficiency in pT [20,40]GeV and it is non-negligible
  - Estimations on Topo-clusters and Particle-flow seeds show they are similar in pT and eta resolution as well as decay mode classification.

- More to do…
  - More inefficiency check:
    - Seed pT influence on PFlow jet efficiency => reach low pT threshold for recover.
    - PFOs => Reco Electron, PFOs will be removed from PFlow jets reco
  - Data/MC comparison for low-pT taus. Modelling
    - $Z \rightarrow \tau_{\mu}\tau_{had}$ tag-and-probe analyses & Other Channels…

Hadronic Tau Substructure

Hadronic Tau Substructure And Precision TES & TER Combined Measurement

- Decay mode identification
  - SubTauID: Hadronic tau decay modes
    - $\tau^- \to \pi^- \nu\ (1p0n)$ $\tau^- \to \pi^- \pi^0 \nu\ (1p1n)$ $\tau^- \to \pi^- \nu + (\geq 2\pi^0)(1pXn)$
    - $\tau^- \to \pi^- \pi^+ \pi^- \nu (3p0n)$ $\tau^- \to \pi^- \pi^+ \pi^- \nu + (\geq 1\pi^0)(3pXn)$
  - Previous classification strategy (R21) Small mismodelling → Further scale factors

- Precision TES & TER measurement
  - Previous: in-situ method, "global"
  - More precise:
    - Different pT region, Different Decay modes, Different Tau ID
  - Get the different resolutions for EM particles and hadronic particles: α_{EM} and α_{Had}

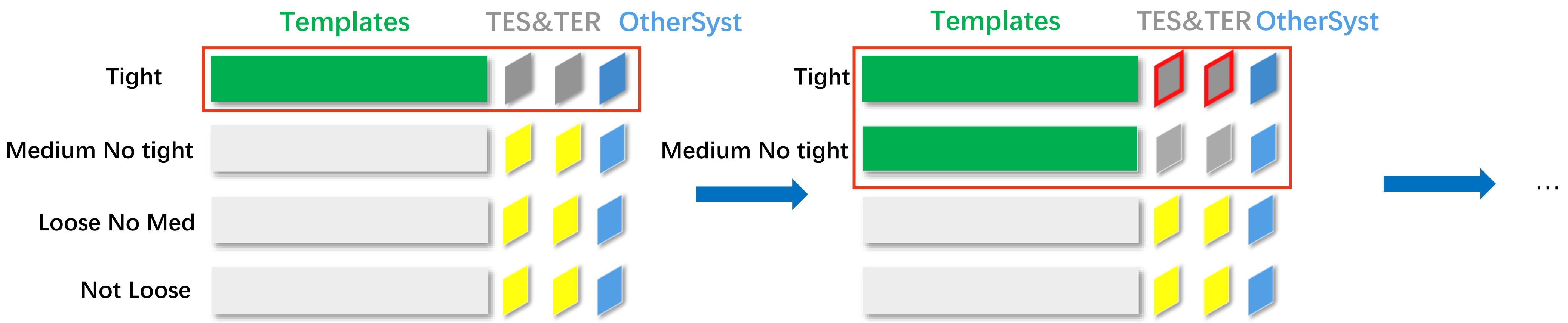
- Combine issue:
  - Decay mode SFs are measured inclusively, but TES are measured in each fit bin.
  - Decay mode SFs are precise to 3 kinds of efficiency fits for each fit bin. While in precision TES measurement are measured in 1 fit
    - $N_{ID}^{Tx-Ry} = N^{Ry} \times \frac{N_{ID}^{Ry}}{N^{Ry}} \times \frac{N_{ID}^{Tx-Ry}}{N_{ID}^{Ry}} = N^{Ry} \times \epsilon_{ID}^{Ry} \times \epsilon_{ID}^{Tx-Ry}$: Reconstruct-, ID-, classification- efficiencies

Strategy

- Measured reco, ID, classification efficiencies By one likelihood fit

$$SF_{Reco} = N^{Ry,Data}/N^{Ry,MC}, \quad SF_{ID}^{Ry} = \epsilon_{ID}^{Ry,Data}/\epsilon_{ID}^{Ry,MC}, \quad SF_{ID}^{Tx-Ry} = \epsilon_{ID}^{Tx-Ry,Data}/\epsilon_{ID}^{Tx-Ry.MC}$$

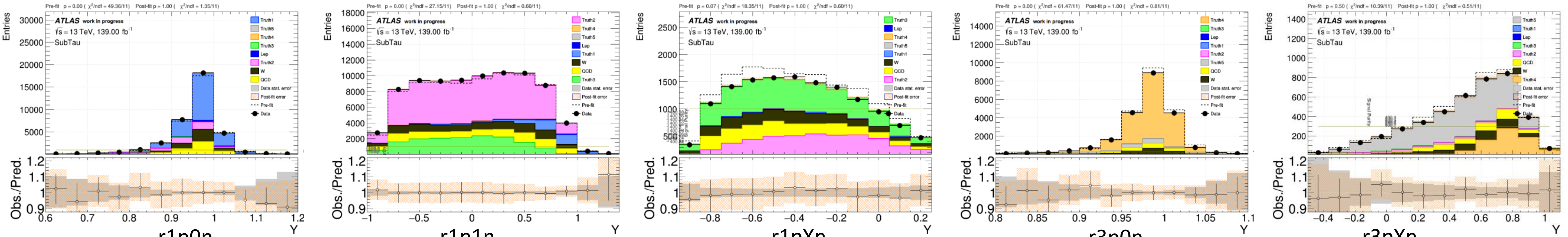
- TES & TER are used as nuisance parameter →get by pull plots
- The fit is binned with Pt, Prongs and tauID. Included all of the systematics.
- A set of recursive fits to get both inclusively ID and binned systematic:



- Until we get the Pre & Post fit ϵ_{ID}^{Tx-Ry} & TES & TER in eachworking point and tauPt bins

Post-fit Result of tight

- The fit result for tight working point tauPt [20,30]
  - High fit quality and fine agreement between the data & post fit template!



r1p0n r1p1n r1pXn r3p0n r3pXn

- Pulls: tight working point tauPt tauPt [20,30] 1 Prong :

TES_EM(%)	TES_Had(%)	TER_EM(%)	TER_Had(%)
$-8.47^{6.97}_{-6.97}\ ^{2.35}_{-2.35}$	$-0.36^{4.36}_{-3.09}\ ^{3.04}_{-3.04}$	$10.34^{6.95}_{-7.72}\ ^{0.64}_{-0.64}$	$-3.54^{5.81}_{-6.23}\ ^{0.39}_{-0.39}$

- The fit will continue for medium working point, Loose …

Results & Outlook

tauPt [20, 30]

tauPt [30, inf]

- The Scale factor: $SF_{ID}^{Tx-Ry} = \frac{\epsilon_{ID;post}^{Tx-Ry}}{\epsilon_{ID;pre}^{Tx-Ry}}$
  - tight working point
  - medium working point
  - loose working point

  - Most of the efficiency are close to 1. Except for SF^{T1-R1} because mis-modling in $1\mathrm{P}x\mathrm{N}$
  - Small errors for most of recons, especially for 3P
  - Medium and Loose WPs have Similar results as tight.

- Future…
  - Check tau pT binning influence on TES fit results. To see if it could reduce TES error.
  - Summarize current results. And use current result for precision TES.

Thanks!

Backups

Samples and pre-selection

- Reconstruction efficiency
  - $\epsilon_{reco} = \frac{N_{Taujets\ matched\ Truth}}{N_{Truth\ Tau}}$

- Samples R22 Ztautau (MC only):
  - Particle-flow seed jets: valid1.361108.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Ztautau.recon.AOD.e5112_s3227_r11986
  - LCTopo seed jets: valid1.361108.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Ztautau.recon.AOD.e5112_s3227_r11978

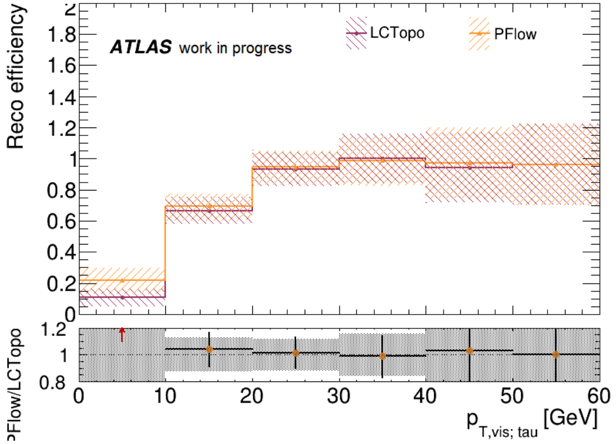
- Pre-selection:
  -

Truth Tau	Reco Tau (Taujet)
Hadronic decaying	Good Track>=1
$p_{T_{vis}} > 10\ GeV$	
$0 < \lvert\eta\rvert < 1.37 \cup 1.52 < \lvert\eta\rvert < 2.5$	
tau-jet matching by $\Delta R < 0.2$	

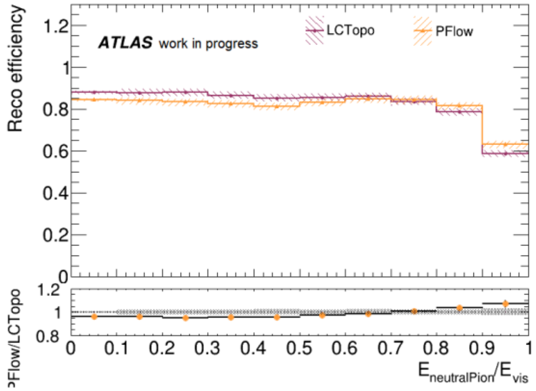
- CMS: performance , algorithm

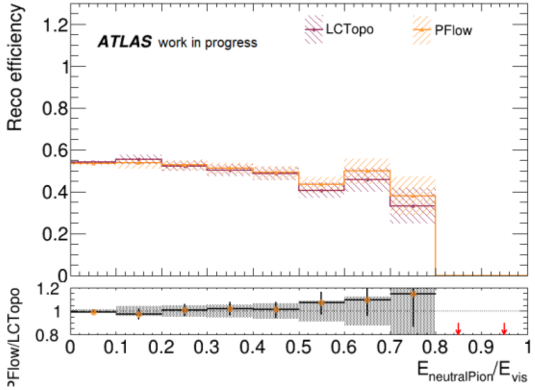
Results

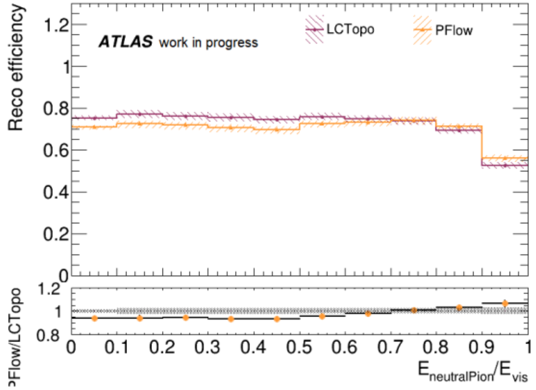
- Reco Electron check
  - $\epsilon_{Ele} = N_{TruthTau\ Matched\ to\ RecoJet+\ RecoEle} / N_{TruthTau\ Matched\ to\ RecoEle}$



- PFLow is better in higher neutral fraction part.

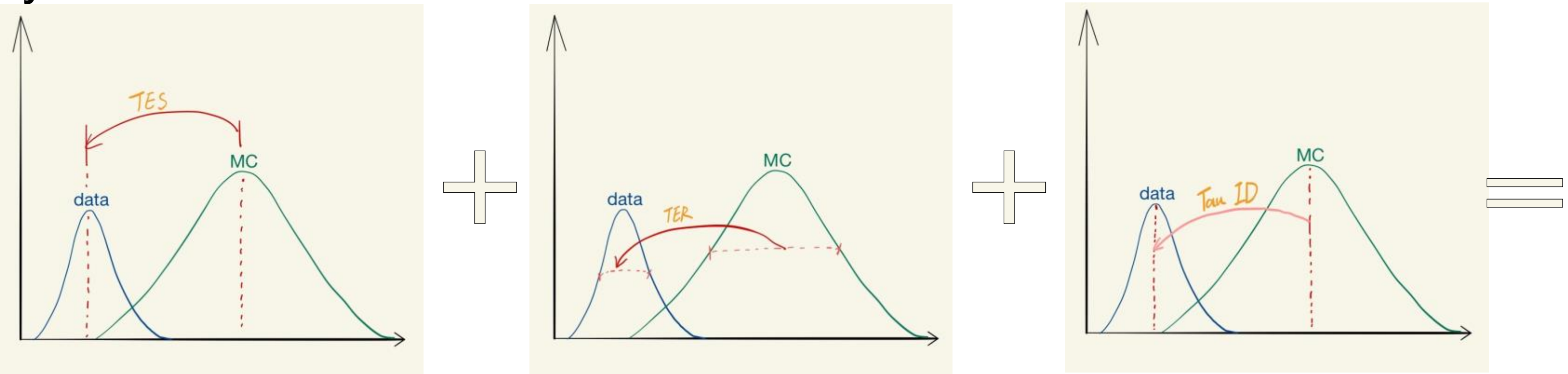






Why combine?

- They have relations with each other…



Agreement!

- All sensitive to the "upsilon variable": $\frac{|\,p_{T,charge} - p_{T,neutral}\,|}{|p_{T\tau}|}$

- $(\sum_{x\in Fixed}^{Tx} N_{MC}^{Tx-Ry} + \sum_{x\in Fit}^{Tx} a_{Tx}\, N_{MC}^{Tx-Ry}) \rightarrow N_{Data}^{Ry}$

Sample & pre-selection

- Samples
  - Data: full Run2-data ($139\ fb^{-1}$)
  - MC samples: $Z \rightarrow \tau\tau$, $Z \rightarrow ll$, Top($t\bar{t}$, single top),$\mathrm{W} \rightarrow l\nu$ and $\mathrm{W} \rightarrow \tau\nu$
  - Samples produced by xTauFramework R21 V10. MVA_TES and RNN tau ID

- Event Selection
  - Target: Select a $Z \rightarrow \tau\tau$ events enriched region. Lep-Had channel

- Base selection
  - $\geq 1PV$, exactly 1 muon, no electrons, no b-tagged jets
  - pass trigger HLT_mu26_ivarmedium or HLT_mu50

- Muon selection
  - offline p_T trigger dependent
  - $|\eta| < 2.5$
  - Medium ID

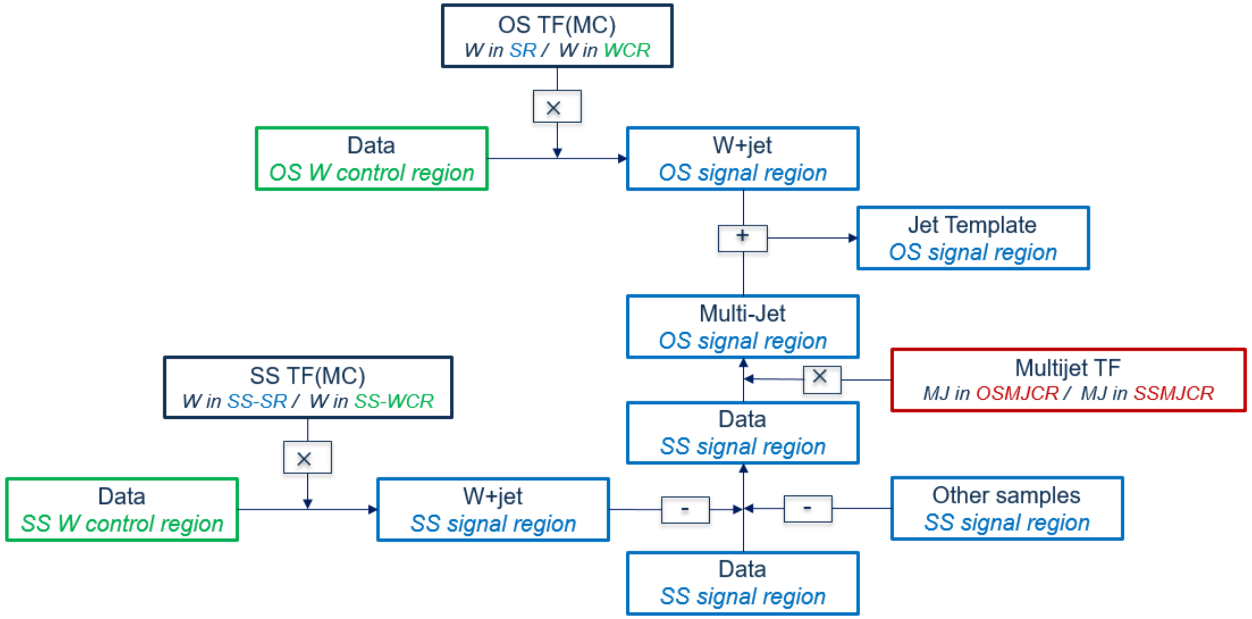
- Tau selection
  - $p_T > 20GeV$, leading tau candidate
  - 1 or 3 tracks, $|q| = 1$
  - Opposite sign (OS) to muon
  - $|\eta_\tau| < 2.47$, excluding $1.37 < |\eta_\tau| < 1.52$
  - $transBDT > 0.005$

Note:

- M_{vis} :The visible mass of the combination of the tau and muon
- transBDT: Jet transverse BDT score. It in applied inside V10

Region definition

- Template construction
  - Tau truth template: $Z \rightarrow \tau\tau$, Top events, where the tau is matched to a truth hadronic tau
  - Lepton template: $Z \rightarrow \tau\tau$, $Z \rightarrow ll$, Top events, where the tau candidate is matched to a truth electron or muon
  - Fake template: W+jets and Multijets. Data driven



The data driven method for fake estimation

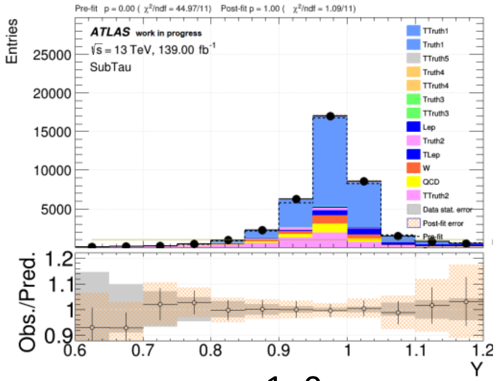
- Region Definitions

- Signal region
  - $M_T^{\mu} < 50GeV$
  - $48 < M_{vis} < 90GeV$
  - $SumCos > 0.15$
  - $|\delta\Phi_{\tau\mu}| > 2.4$
  - muon_FCTightTrackOnly_FixedRad
- W+jets control region
  - $E_T^{miss} > 20GeV$
  - $48 < M_{vis} < 90GeV$
  - $M_T^{\mu} > 60GeV$
  - $SumCos < 0$
  - muon_FCTightTrackOnly_FixedRad
- Multi-jets control region
  - invert isolation
  - $M_T^{\mu} < 50GeV$
  - $48 < M_{vis} < 90GeV$
  - $SumCos > 0.15$
  - $|\delta\Phi_{\tau\mu}| > 2.4$

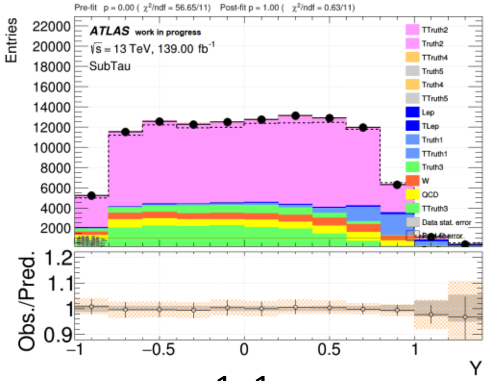
$SumCos = cos(\Delta\phi(e, E_T^{miss})) + cos(\Delta\phi(\mu, E_T^{miss}))$

Results

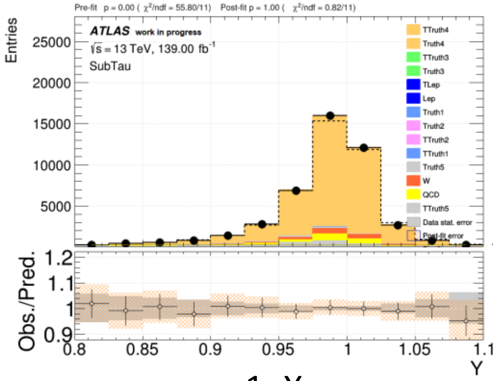
- The fit result for medium working point tauPt [30,inf]
  - High fit quality and fine agreement between the data & post fit template!



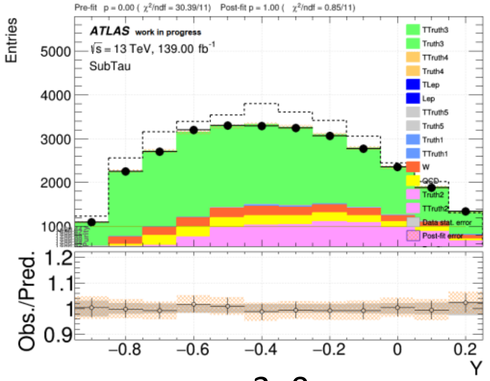
r1p0n



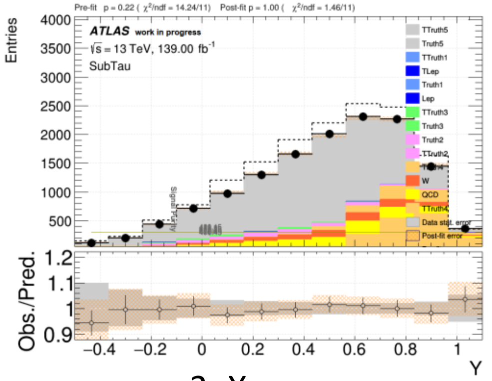
r1p1n



r1pXn



r3p0n



r3pXn

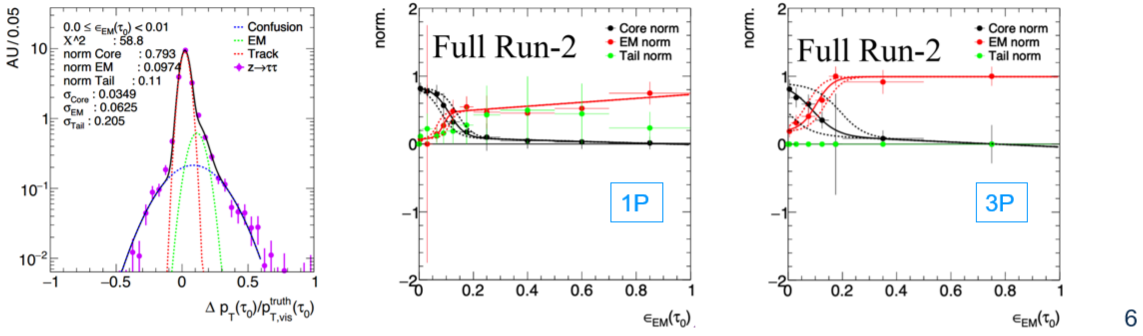
- Pulls: medium working point tauPt tauPt [30,inf] 1 Prong
  - ± 1 sigma of the fermi transition + largest shift on the fermi transition uncertainty

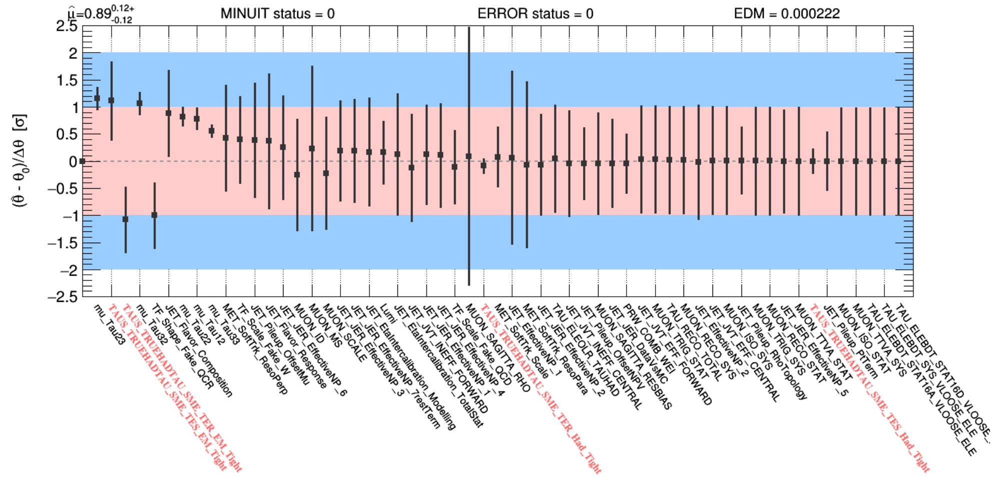
TES_EM(%)	TES_Had(%)	TER_EM(%)	TER_Had(%)
$4.18^{9.14}_{-9.14}\ ^{1.68}_{-1.68}$	$10.49^{8.93}_{-8.93}\ ^{4.85}_{-4.85}$	$7.82^{8.33}_{-8.33}\ ^{5.83}_{-5.83}$	$5.50^{9.15}_{-9.15}\ ^{0.99}_{-0.99}$

TES parametrization

Done By Gang Zhang

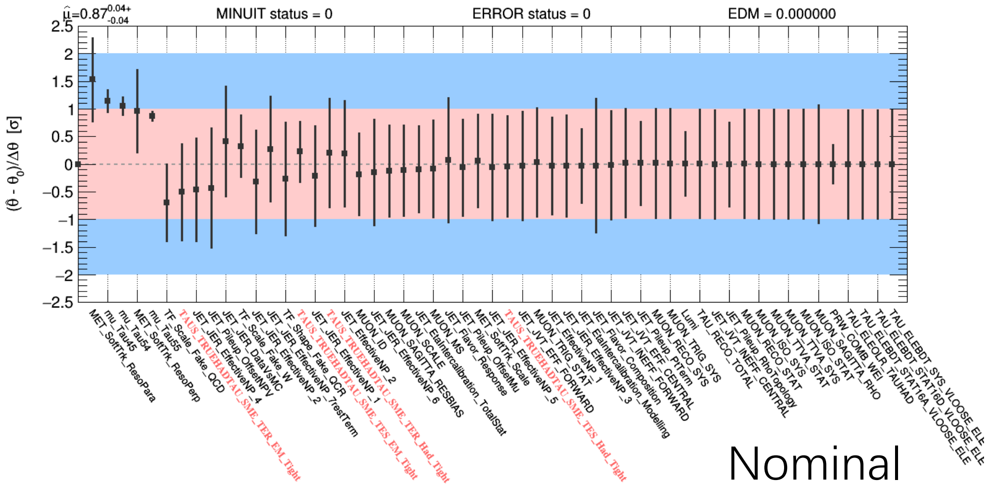
- Fitted by a combination of liner function and a fermi transition
- The uncertainty of this fit is also considered by $\pm$ 1 sigma of the fermi transition
- When doing the fit, two additional fit with this function up/dw will also do. The final TES uncertainty will add an additional uncertainty about the largest shift on the fermi transition uncertainty



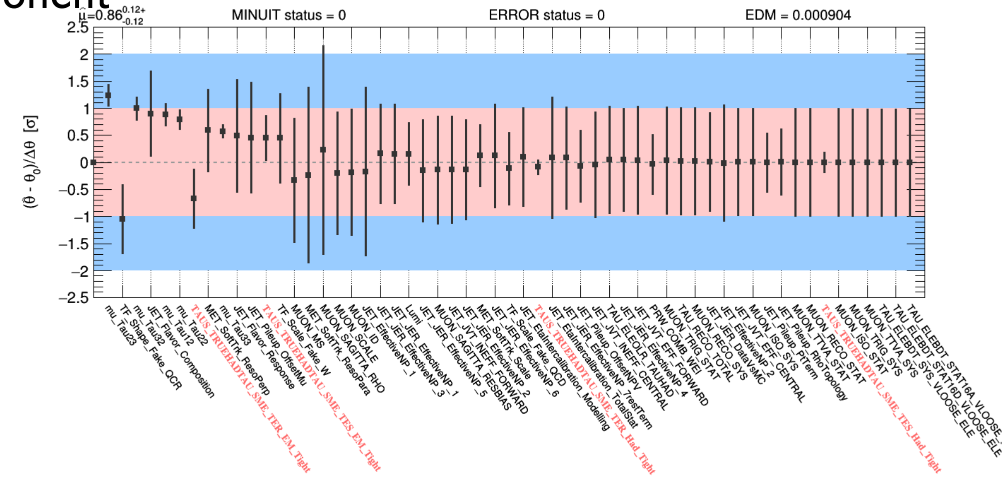


FermiUp

- Different resolutions for EM particles or hadronic like particle: α_{EM} , α_{Had}
- Make a shift on p_T and get the shifted p_T of each EM/Had TES and TER component



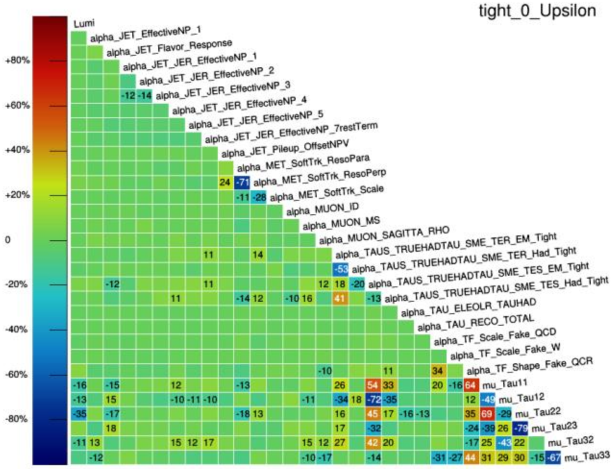
Nominal

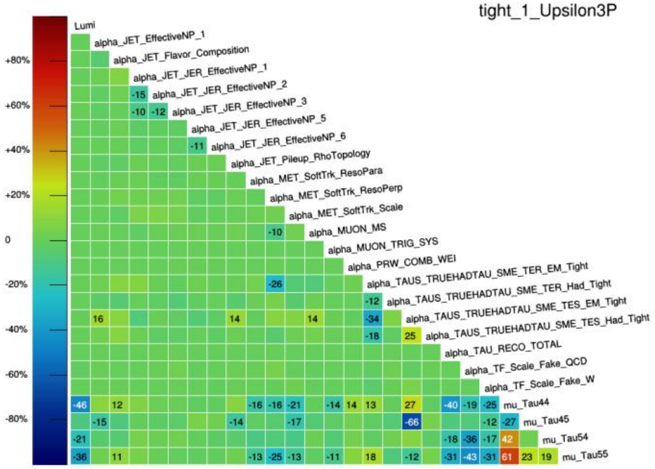


FermiDw

Results

- Systematic tight working point tauPt [20,30] 1Prong





- Pre & Post fit ϵ_{ID}^{Tx-Ry} tight working point tauPt [20,30]