

Machine learning in $\mu + \tau$ channel: status update

Merijn van de Klundert

June 23, 2019



Machine learning in $\mu + \tau$ channel: status update

Reviewed Data-MC agreement with lower pT cutoff 20 GeV;

- Can compare jets graphs to AN. Shapes look in order
- Jet distributions: rather DATA higher than AN plots, while MC appears correct!

Latest new updates (with pt cut 30 GeV):

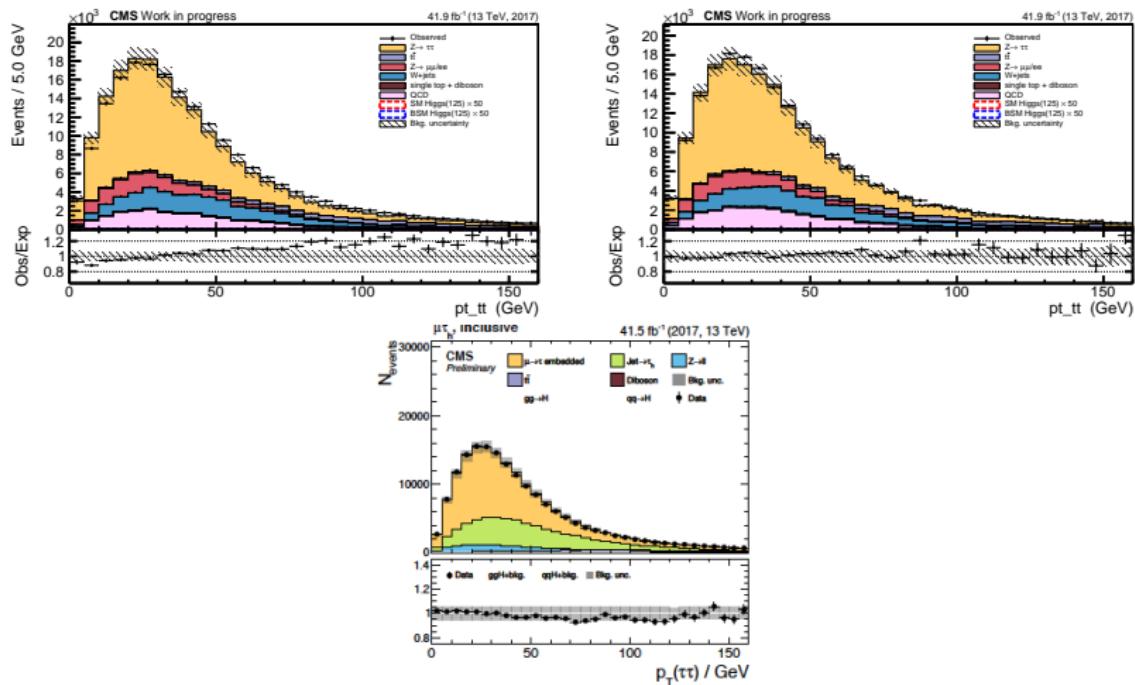
- Corrected the ZptReweigh (the boundaries required attention)
- Updated MET from ptmet to pfmetcorr. NOTE: this also enters the recoil calculation!

Overall 18 distributions

- Review only plots that substantially changed
- LEFT: PREVIOUS RESULT. RIGHT: UPDATED RESULT

pt tau pair

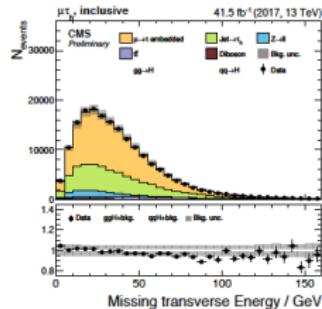
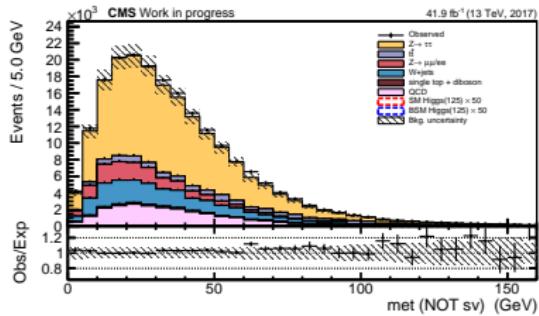
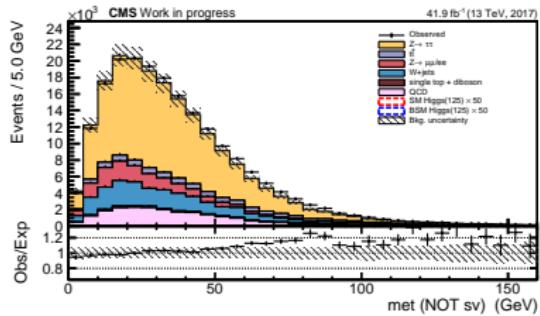
- Updating zpt reweighting to 2017 value solved the discrepancy!



All lepton observables modelled well now it seems

MET

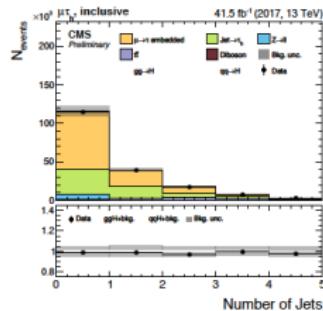
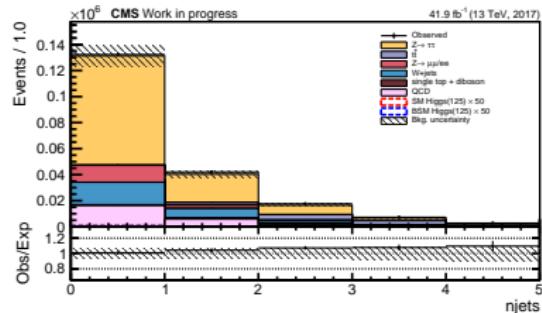
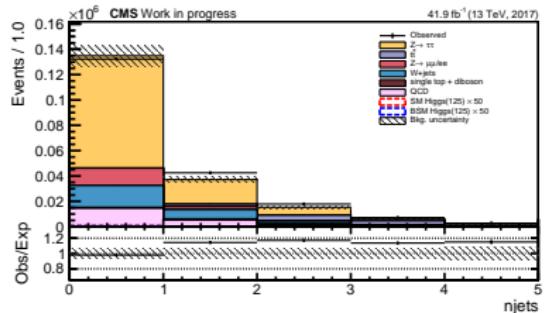
- Updating pfmet to pfmetcorr. restored data/MC agreement!



Jet multiplicity

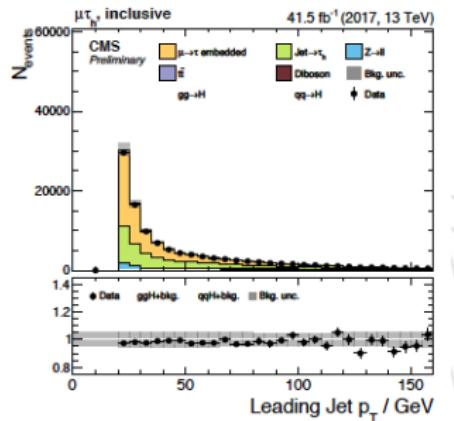
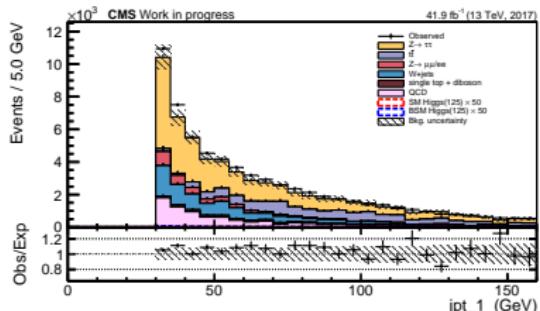
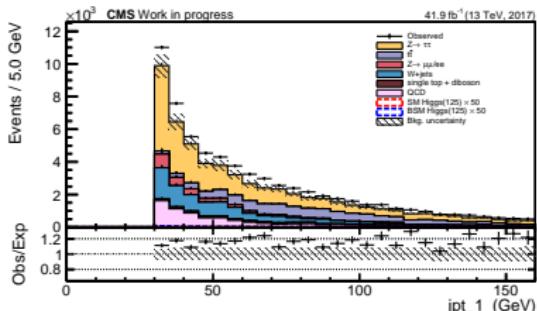
Jet multiplicity

- Highly improved, within uncertainties. Consequence of recoil corrections?



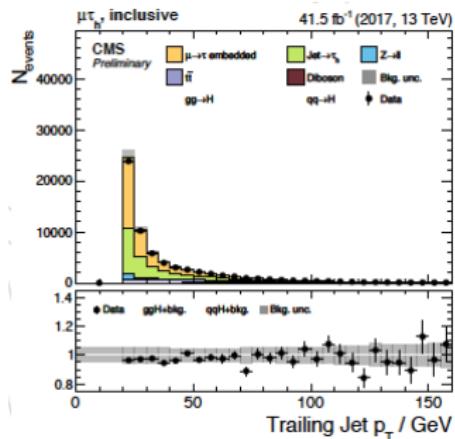
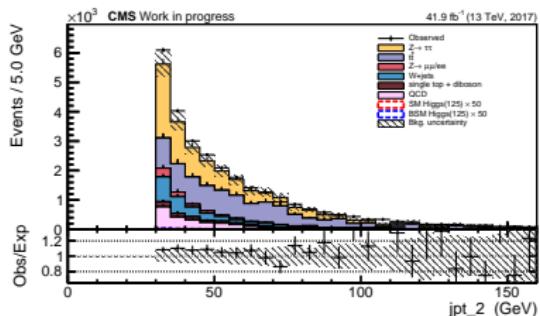
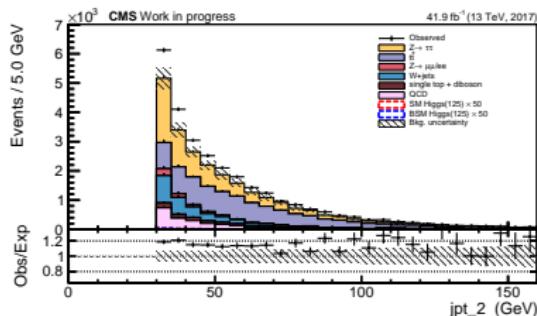
Leading jet

- data/MC agreement improved. Interesting to see if embedded and fake factor improves first bins



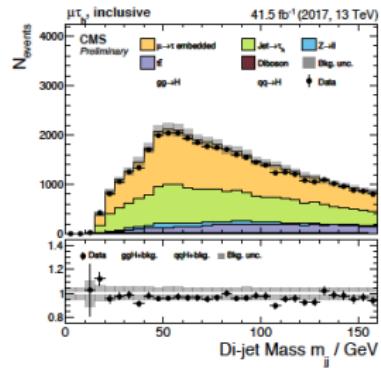
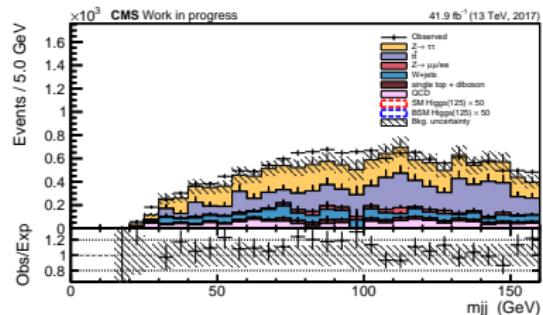
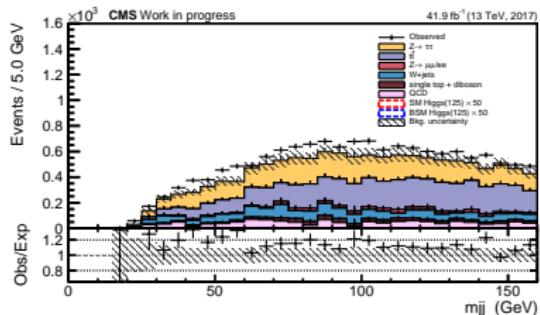
Subleading jet

- data/MC agreement improved. Still some trend for low-pt? Interesting to see if embedded and fake factor improves first bins



jet mass

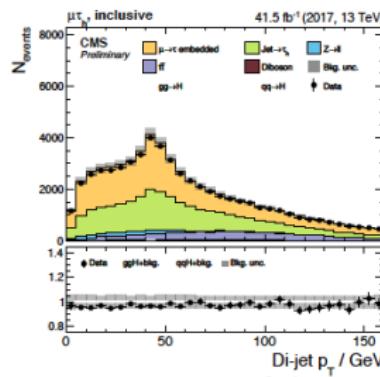
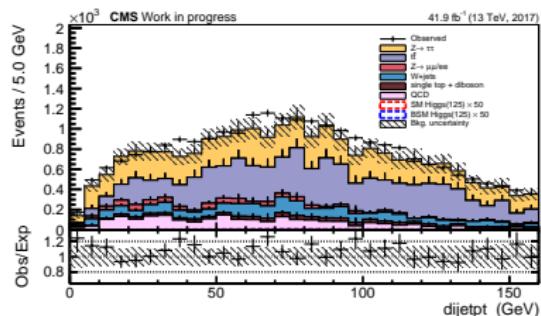
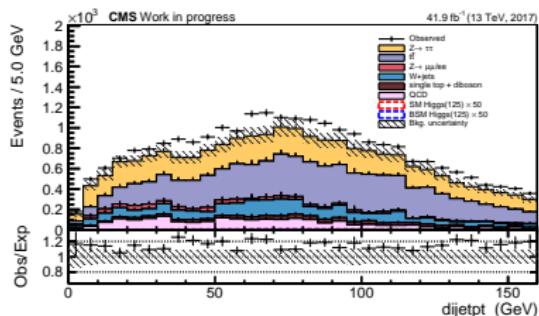
- Systematic offset decreased. Still improvements possible..



Dijet pT spectrum

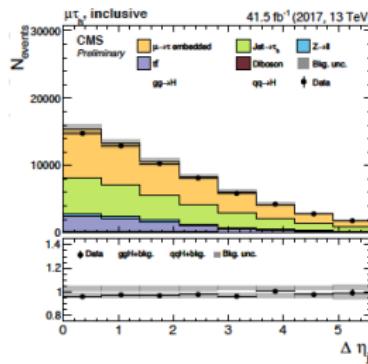
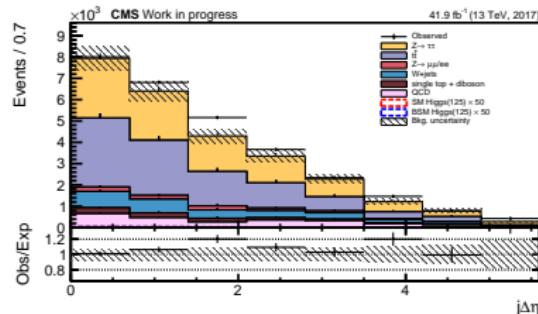
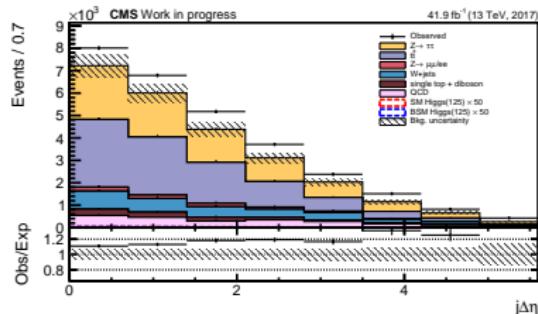
Dijet pT spectrum

- Systematic offset gone, sensible agreement



jet $\Delta\eta$

- Systematic offset gne/highly improved
- Distribution not yet perfect. Here embedded plus fake factor may improve things further



Next steps

Data/MC

- Data/MC agreement appears in order, certainly good enough to proceed. Remaining disagreement could be MC normalisation, or resolved by embedded
- Await for fake factors plus embedded. Need to ask Vienna for certain parameters..
- Once ready, proceed also with Goodness of Fit (GOF) test
- Discuss systematics incorporation with Valeria, plus these anti-isolation cut files
- Produce low-mass DY MC still at some point, not highest priority
- Backup: other observables for latest NTuples

Classification

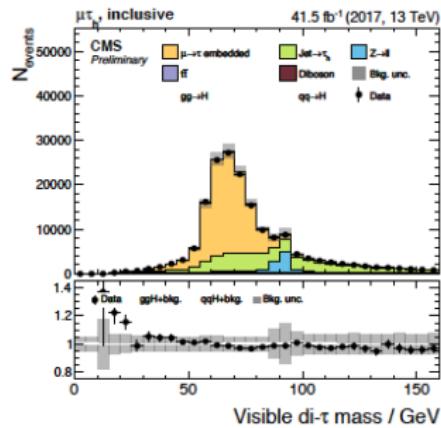
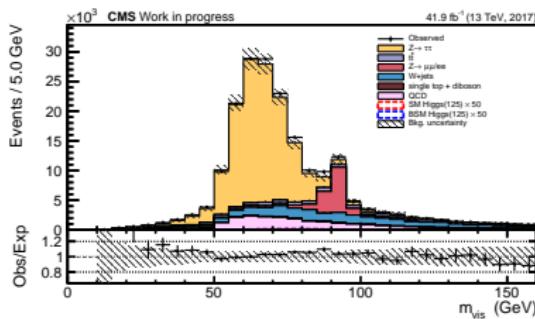
- Categorisation: first results not yet encouraging, certain categories stay empty
- Review the input distributions to DNN, some mismatch?
- Once categorisation working, proceed with producing data cards
 - ▶ Respect conventions from Imperial! Combine template set up already?
 - ▶ Not forget to calculate weights with `calc_weights.py`(!)

Overall

- Seems that data/MC agreement is resolved
- Aim now for inclusion fake factors and embedded

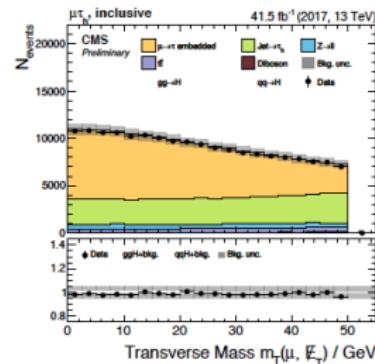
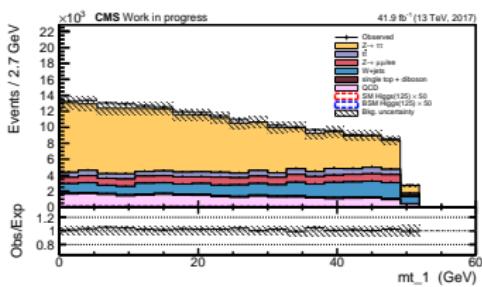
Mvis: sensible data/MC agreement

- Still need to add the low-mass DY
- Shape peak improved with iteration 2. Peak value: reasonably consistent now with AN



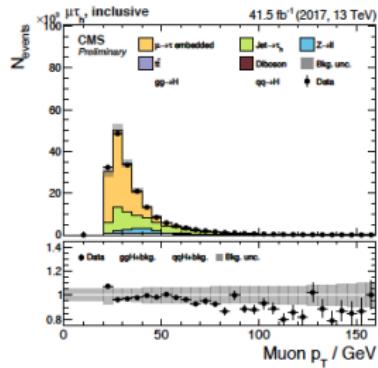
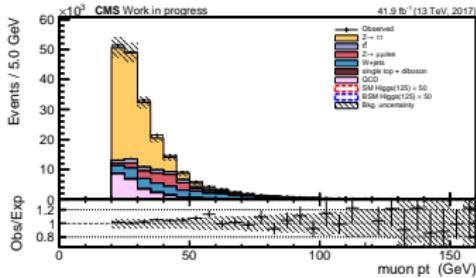
mt muon: good data/MC agreement

- Higher end of spectrum could be worth further reviewing..
- Value AN slightly lower than our value

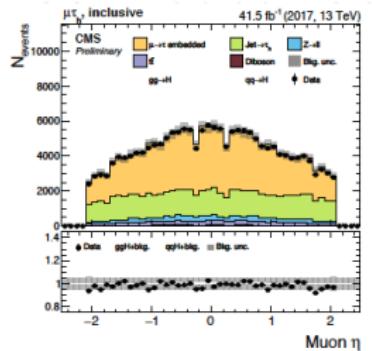
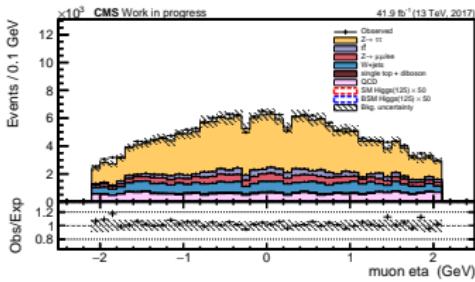


pt and muon: reasonable data/MC agreement

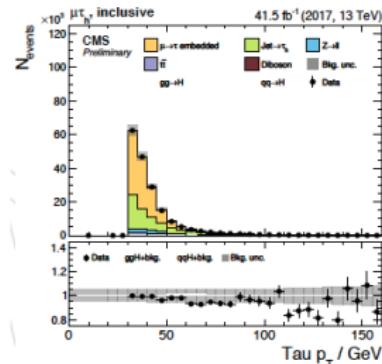
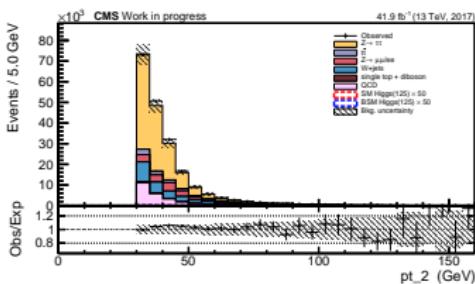
- AN more sharply peaked. Integral first 3 bins does not match with AN



pt and eta muon: reasonable data/MC agreement. Match with AN



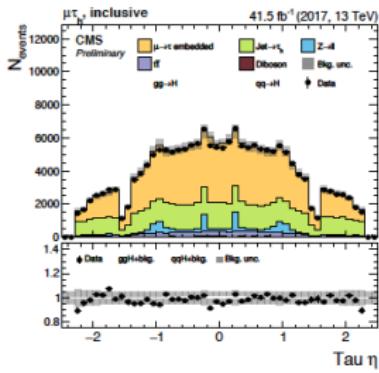
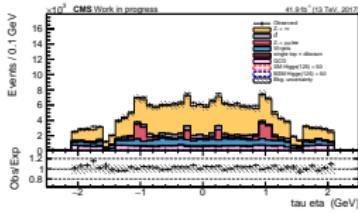
pt and eta tau: reasonable data/MC agreement. Peak AN slightly lower



Lepton observables seem modelled reasonably well

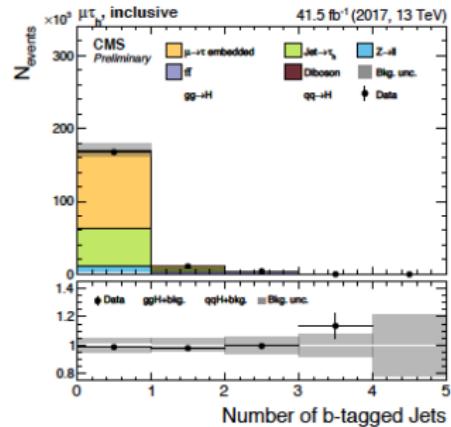
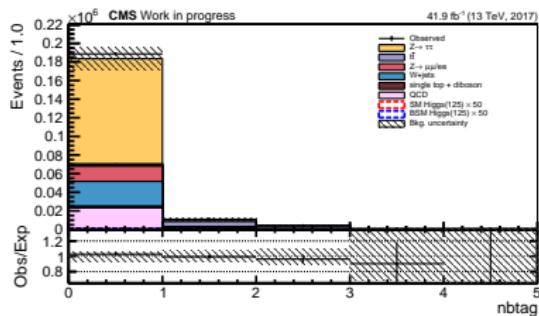
Full comparison to AN plots once binning adjusted..!!

eta tau: reasonable data/MC agreement



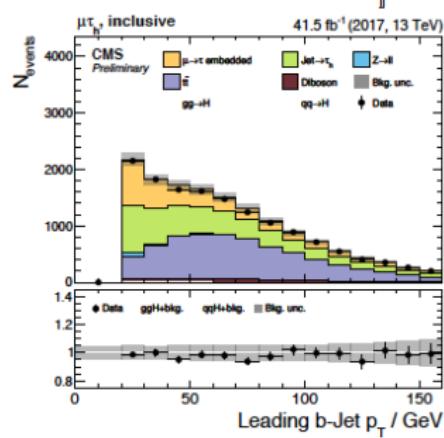
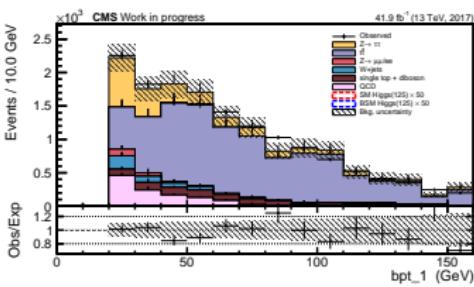
Number BTags

Highly improved! Good data/MC agreement. Good consistency with AN



pT spectra leading b-jet

Highly improved, good data/MC agreement, norm matches AN



pT spectra trailing b-jets

Highly improved, good data/MC agreement, norm matches AN

- Overall, for bjets highly improved

