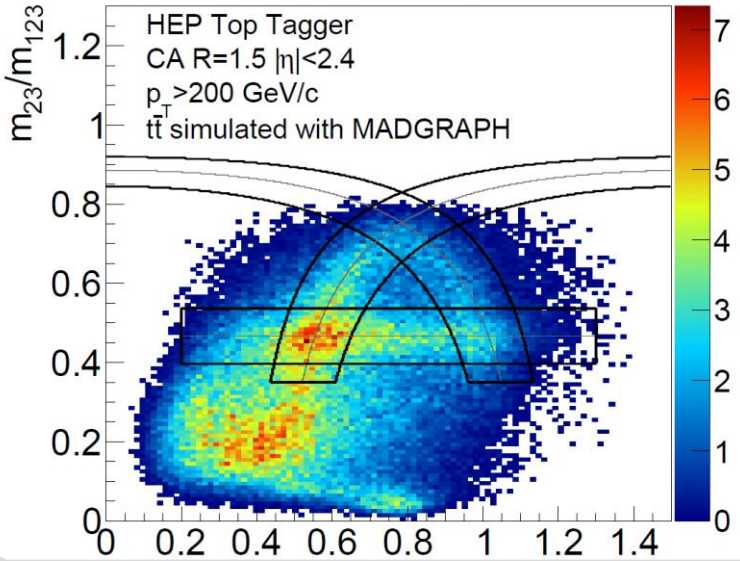
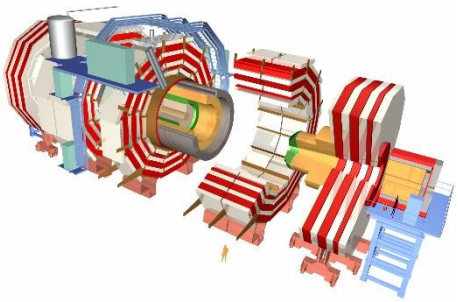
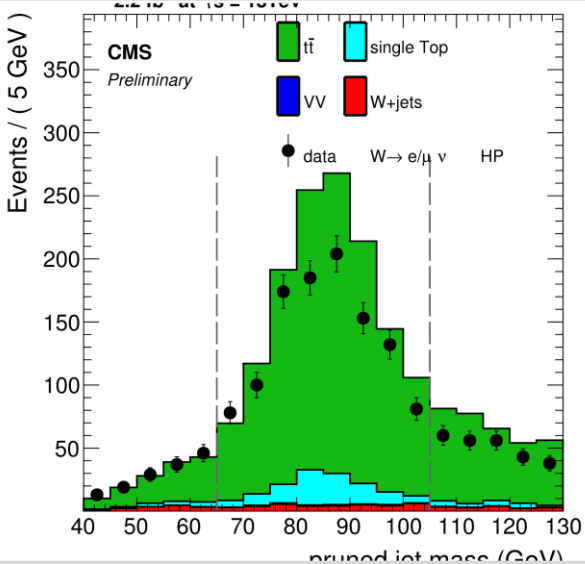
Jet Substructure Tools in CMS
W/Z/H/t

ATLAS/CMS MC workshop
11.01.2016

Matthias Mozer (KIT)

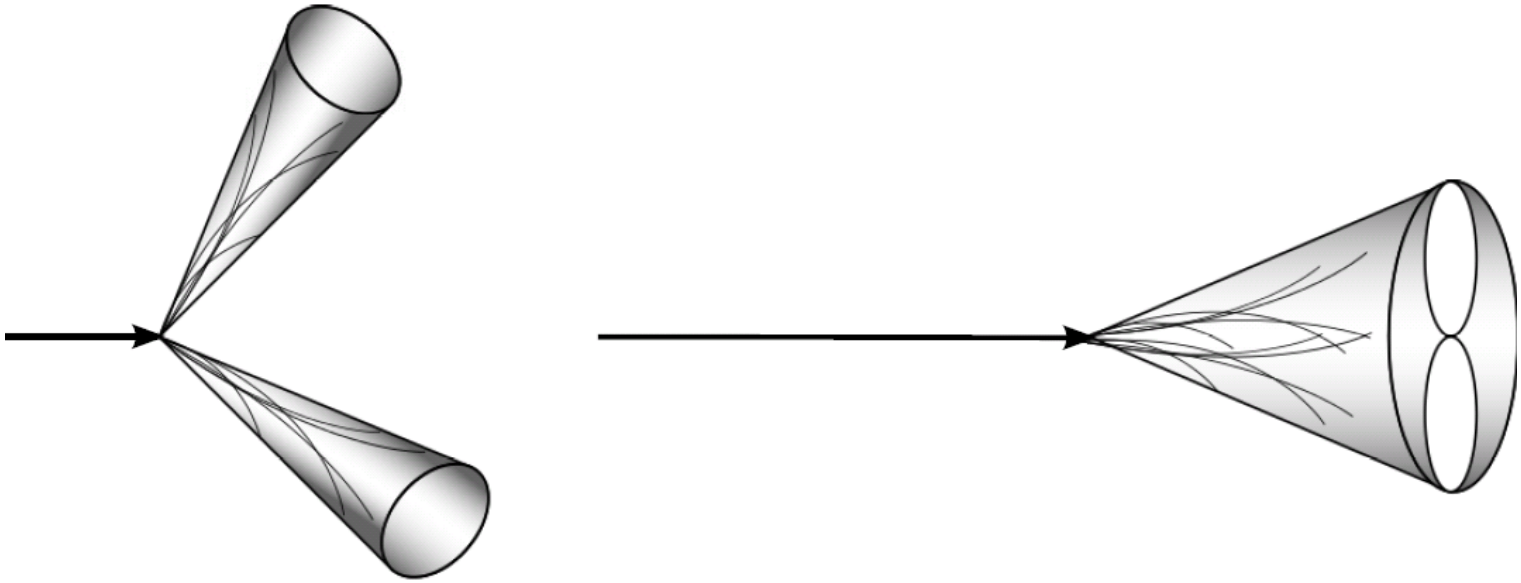






Jet Substructure

- W/Z => two-prong structure
  => jet mass
- H/Z => add b-tag
- t => three-prong structure
  => W mass

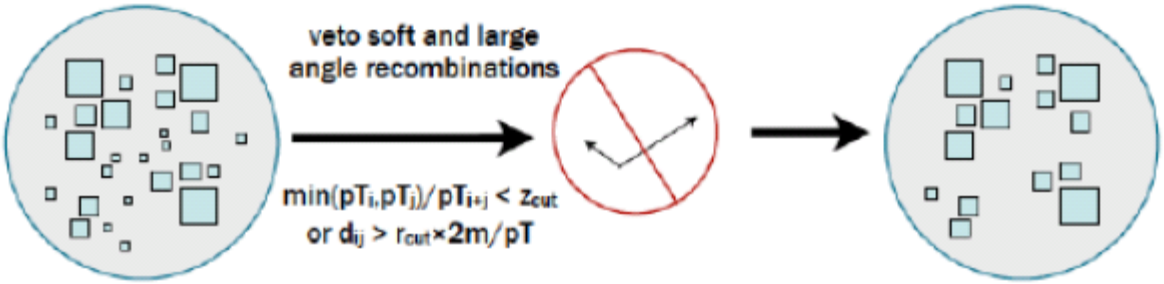


W/Z Run I Reminder

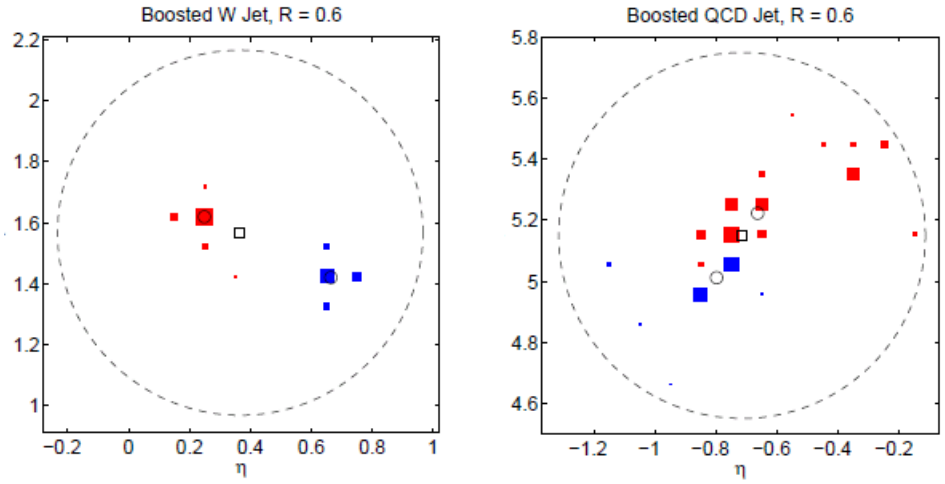
- pruned mass

$$z = \frac{min\left(p_{T,i}, p_{T,j}\right)}{p_{T,JET}} > 0.1$$

$$\Delta R < 0.5 \frac{M_{JET}}{p_{T,JET}}$$



- n-subjettiness

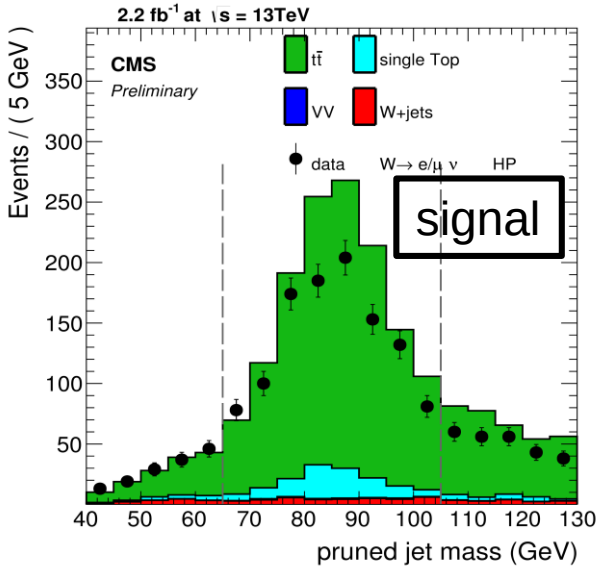


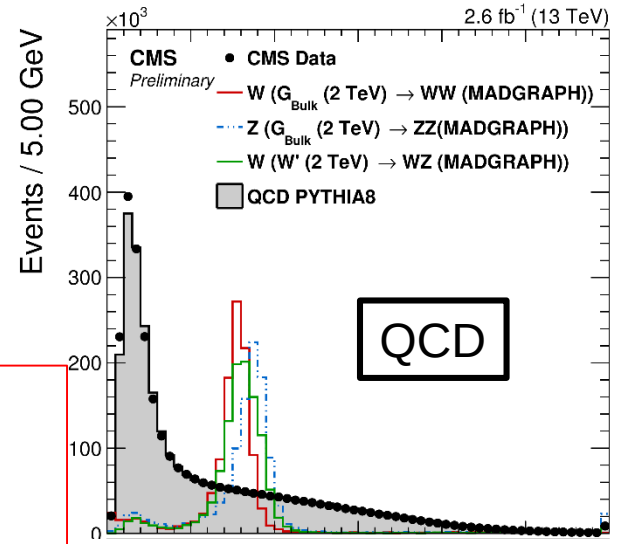
JHEP 1103:015,2011

Also used for Run II early analysis

Run I vs Run II

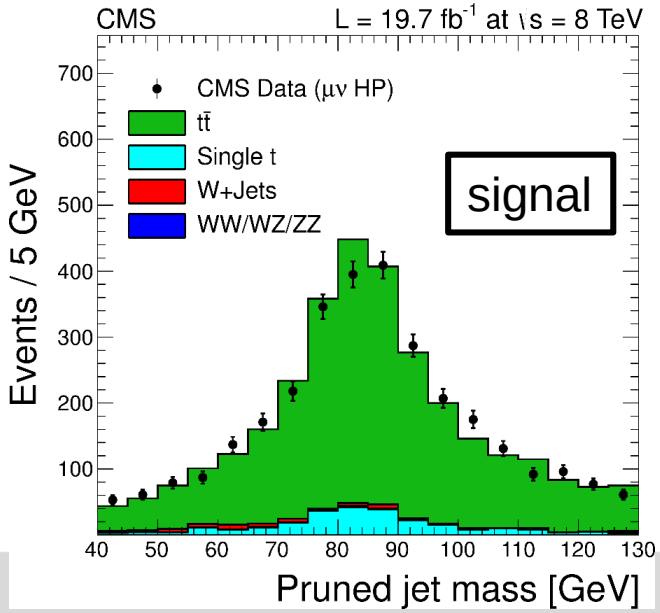
Run II

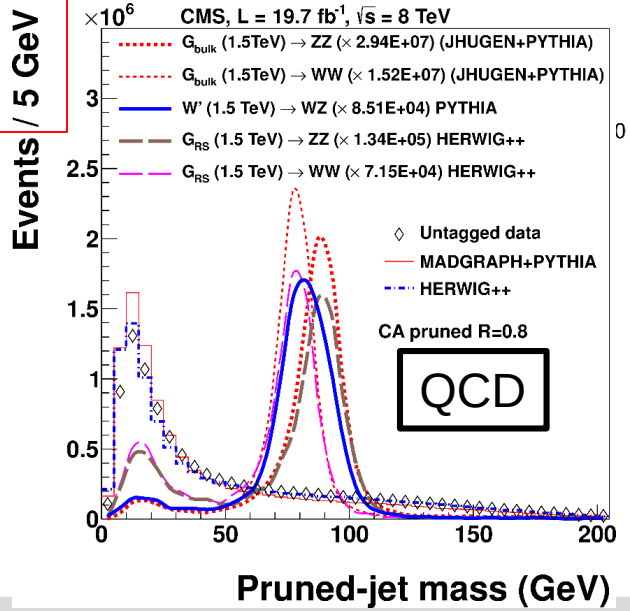




updated:
- reco
- simulation
- Pythia version

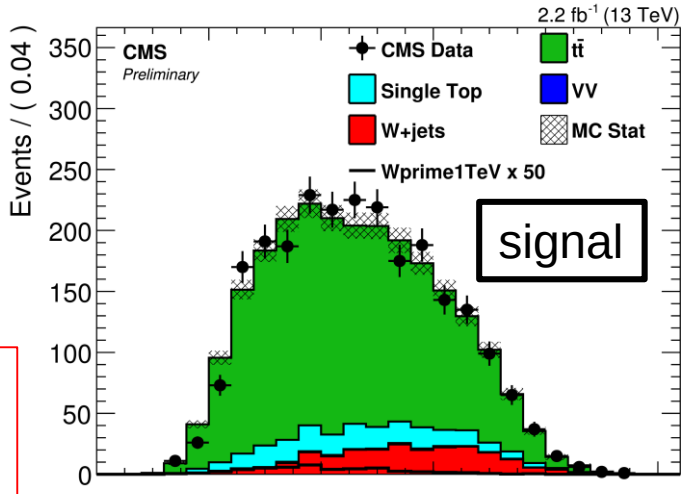
Run I

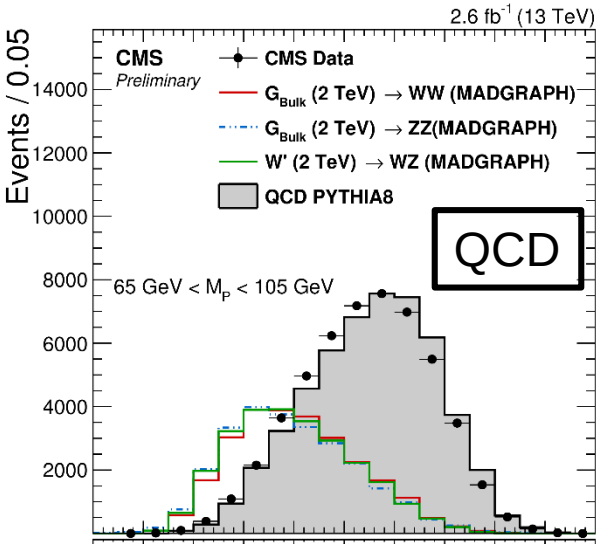




Run I vs Run II

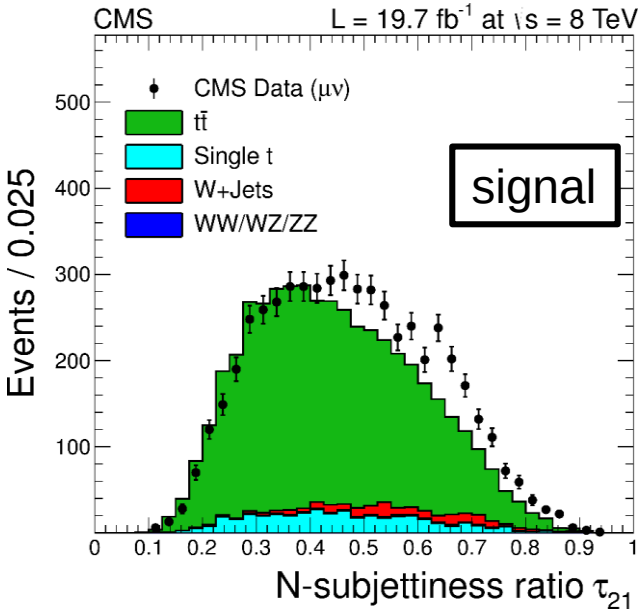
Run II

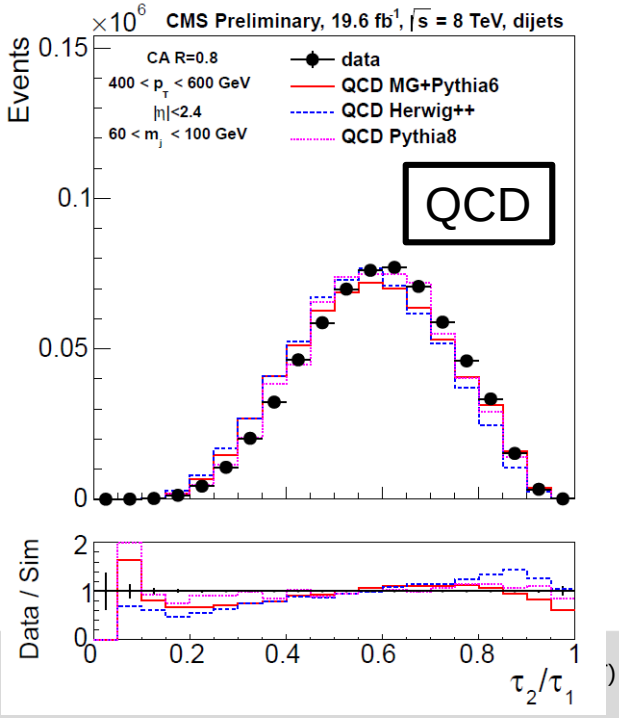




updated:
- reco
- simulation
- Pythia version

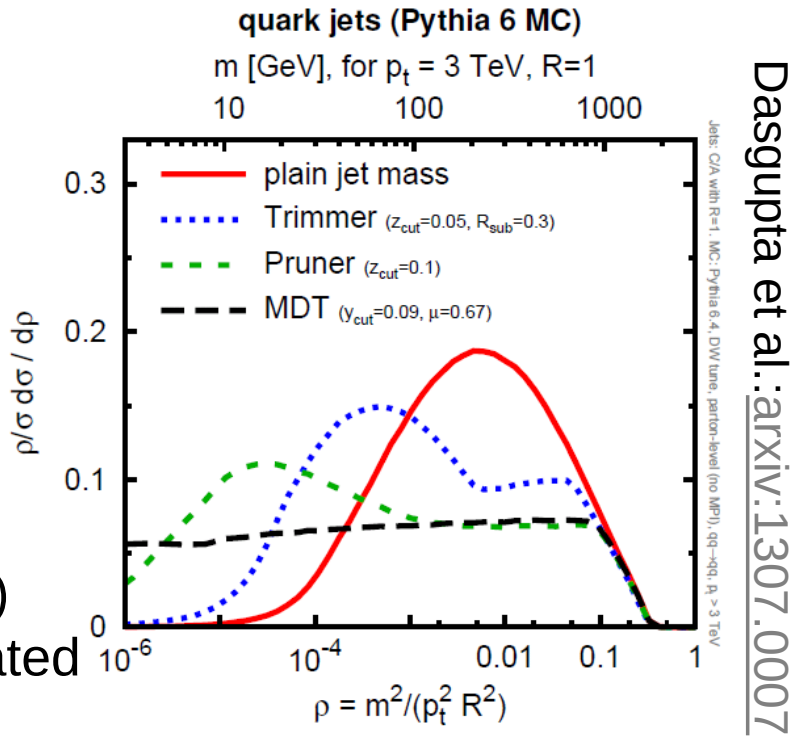
Run I



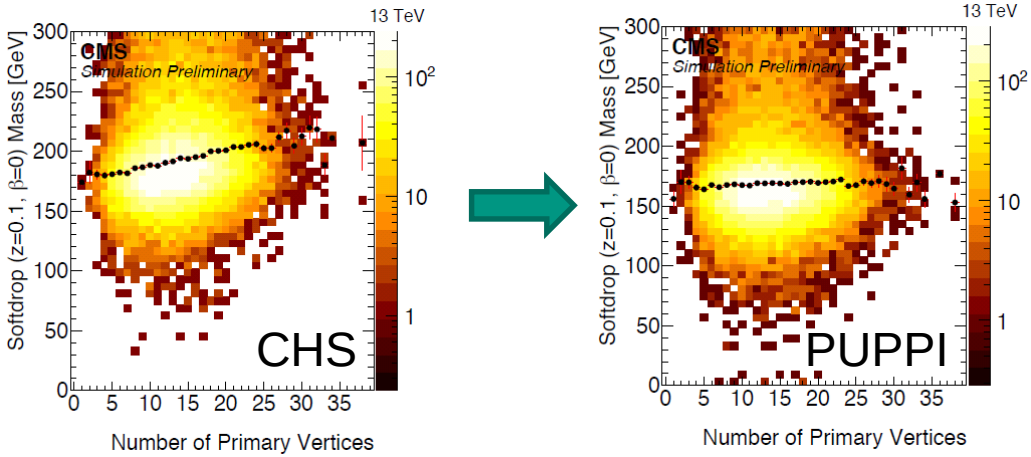


Run II plans

- pruned mass
  => PU dependence
  => scaling behaviour
- Plan for Run II
  => switch to mMDT
  (softdrop, β=0)
  => better scaling behaviour)
  => requires more sophisticated PU treatment (PUPPI)
  => studies ongoing

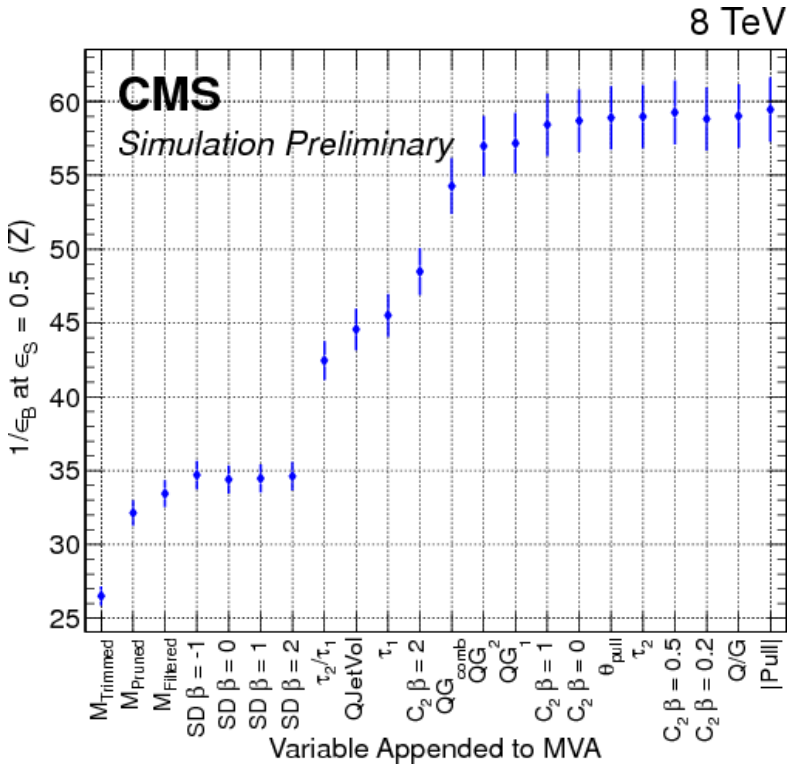


Dasgupta et al.:arxiv:1307.0007



Run II plans

- N-subjettiness used for early results
  => consistently good performance
- Looked at more fancy variables:
  => energy correlation functions
  => subjet pull
  => …
- No single better performer
  than N-subj. + mas
- Significant gains from MVA
  => will need more commissioning work



Subjet btags: H/Z

- Reminder Run I

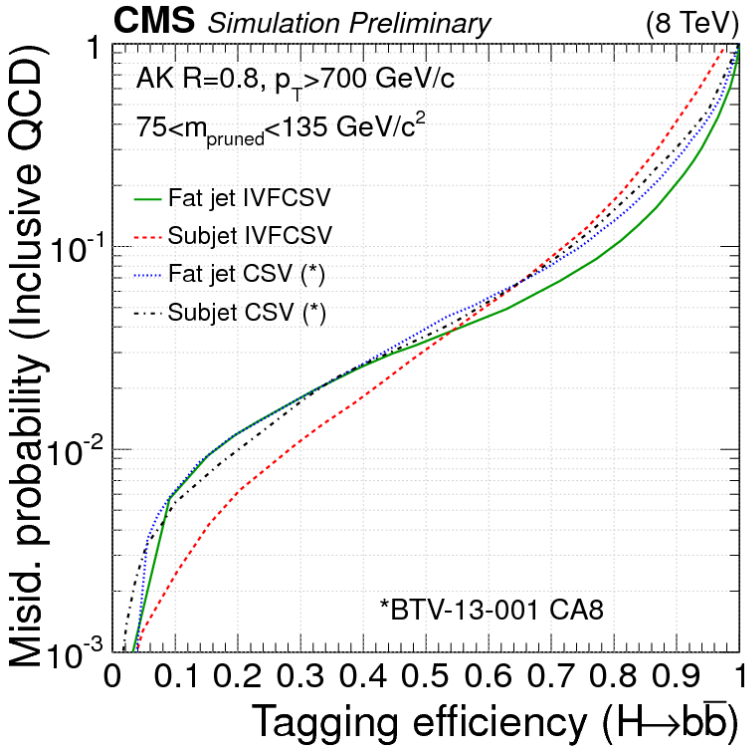
=> used standard b-tagger
=> mixture of subjet btags and fat-jet b-tag
pt dependence of efficiency

=> subjet b-tag tricky for very high pt

=> Fatjet b-tag prone to fakes

=> difficult kinematic dependences

Easy to develop but difficult to check
=> quick results
=> increased systematics

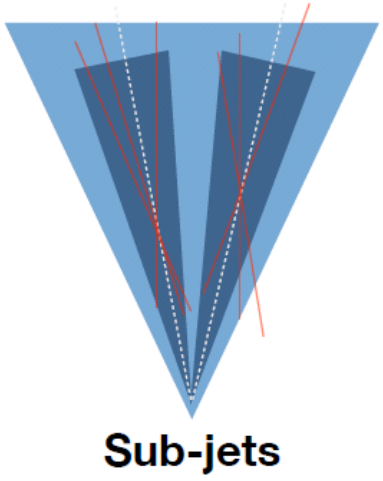


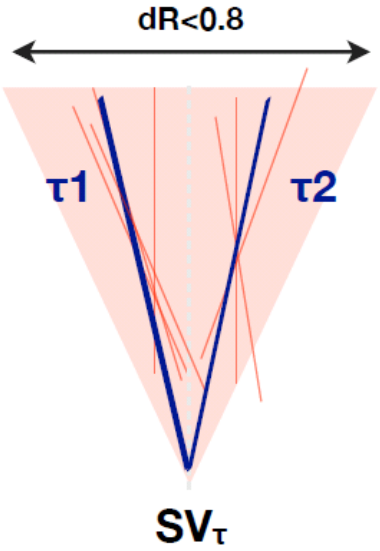
Run II plans for subjet b-tag

- Complete revamp
  => based on updated „normal“ b-tagger

- Major points for optimization: performance independent of:
  => jet mass
  => jet radius
  => jet pt

- Studies ongoing
  => no public results yet



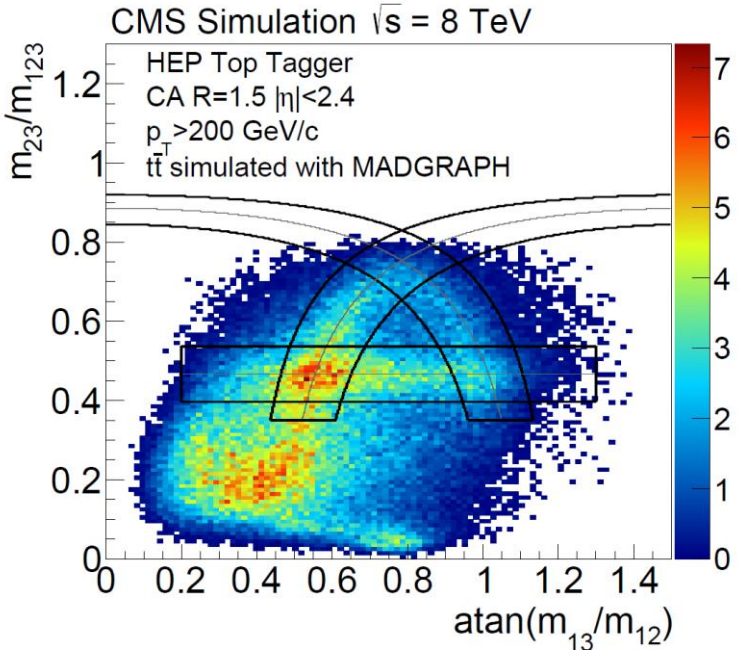


Top-tag: Reminder Run I

- HEP-Top-tagger
  => very wide jet (R=1.5)
  => complex cleaning procedure
  => reconstruct top and W mass
  => best performance at lower pt

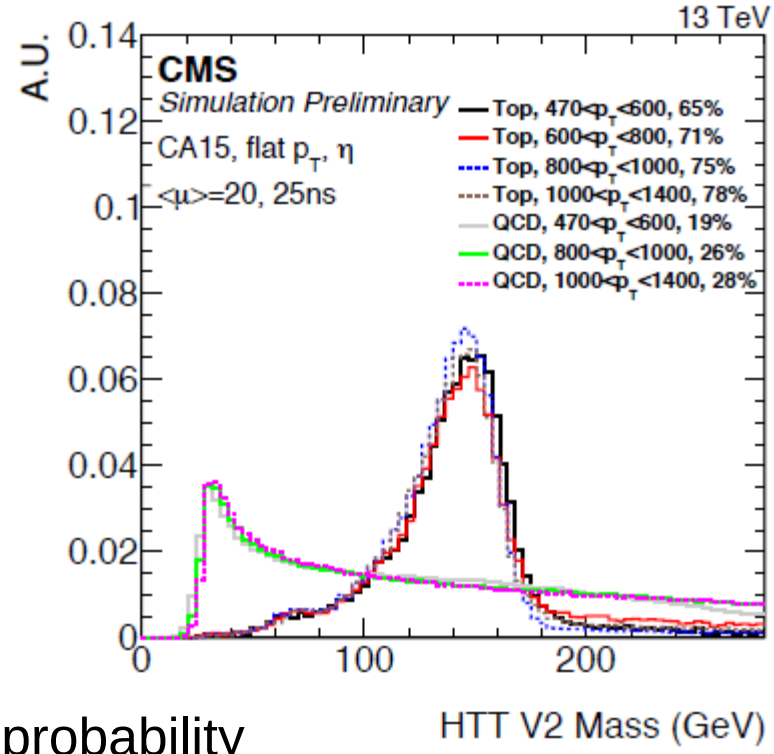
- CMS-Top-tagger
  => relatively simple 3-subjet finder
  => reconstruct top and W mass
  => best performance at high pt

- Both algorithms combined with
  => subjet btag
  => n-subjettiness

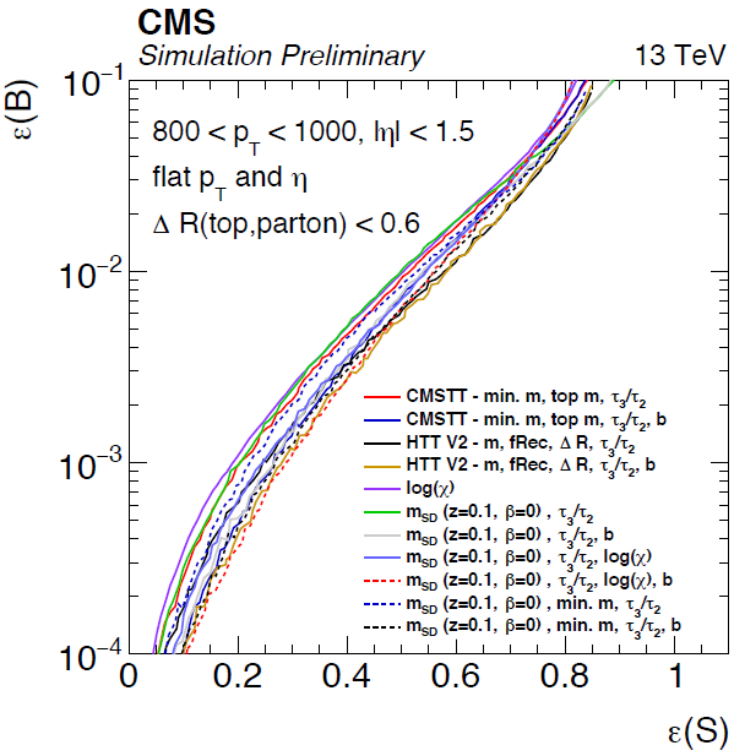


Run II developments

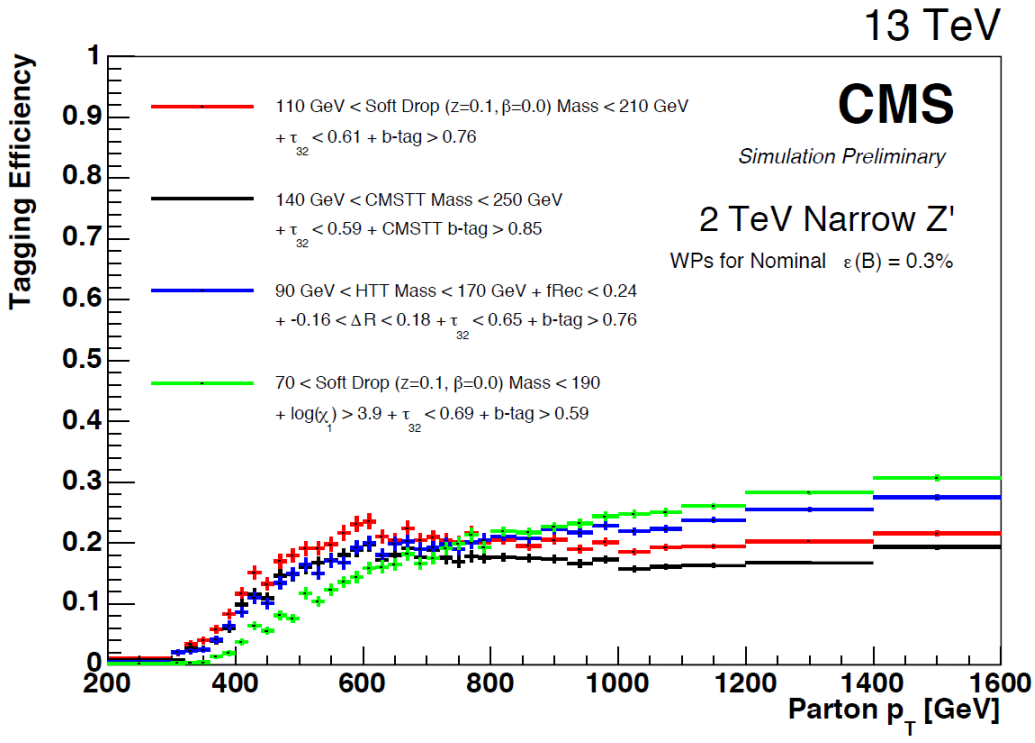
- CMS-Top-tagger (updated)
  => a simple baseline
  => mMDT mass + Nsubjettiness
- HEP-Top-tagger improved:
  => variable radius improves
  high pt behaviour
- New tagger studied:
  Shower deconstruction
  => recluster into microjets
  => use likelihood discriminant from
  shower simulation to determine top-probability
  => excellent performance
  => fake-rate + efficiency very pt dependent, tricky to us in practice
- trying to keep variables similar to W/Z/H-tags at high pt
  => mMDT mass, Nsubjettiness, (looking at PUPPI)



Run II Performance expectations



Best performance with combined criteria



pt-dependence of efficiency (and fake-rate) troublesome for shower deconstruction

Conclusion

- Many new developments for Run II
  => softdrop + PUPPI jet masses
  => improved subjet btag
  => updated top-taggers

- Validation on data takes time:
  => December results still use legacy methods

- Trying to have results with improved taggers soon