Reconstruction and identification of hadronic objects with CMS

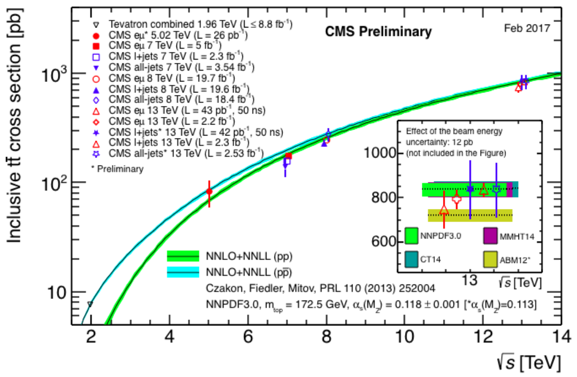
Mauro Verzetti, University of Rochester
on behalf of the CMS Collaboration
EPS2017 - Venice

What do we identify?

- $\tau \rightarrow \nu$+hadrons - TAU-16-002
- Jet flavour:
  - b-jets - DP-2017-013
  - c-jets - BTV-16-001
  - quarks vs. gluons - CMS-JME-16-003
- Boosted objects:
  - t → bqq - CMS-JME-16-003
  - W/Z → qq - CMS-JME-16-003
  - X → bb (not discussed in this talk) BTV-15-002

Why? - from our spokesman LHCP highlights

Top Pair Cross Sections



CMS: 835 ± 33 pb
Theory: 816 ± 42 pb

Top pair rate is > 10 Hz, enabling us to address much more precise questions

- Single and double differential cross sections
- Rare (FCNC) decays
- CP violation (a beginning)
- Width and more complex methods for measuring the mass

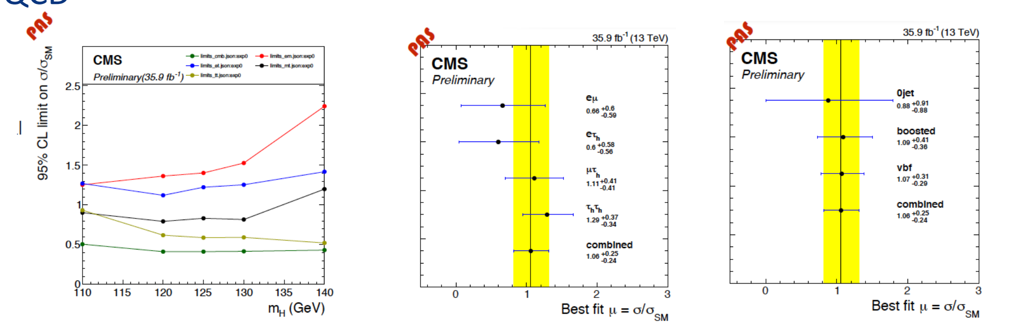
Factory	Quark	Cross Section (nb)	Luminosity ($cm^{-2}s^{-1}$)
B (KEKb)	Bottom	1.15 (Υ(4S))	2.11×10^{34}
LHC	Top	0.82 (incl t-t)	1.51×10^{34}

Top pair production at 13 TeV CM energy is mainly (80%) produced by gluons, providing important information on the gluon distribution at relatively high x, up to ~0.25

Higgs → $\tau^+\tau^-$

HIG-16-043

- Four decay topologies for $\tau^+\tau^-$: $e\mu$, $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$
- Three production modes: 0-jet (gg), VBF, boosted (additional objects)
- Irreducible sources of systematics: W+jets, DY Z/γ → ll,ττ, t-tbar, QCD



$\tau^+\tau^-$, $\mu\tau$ most sensitive

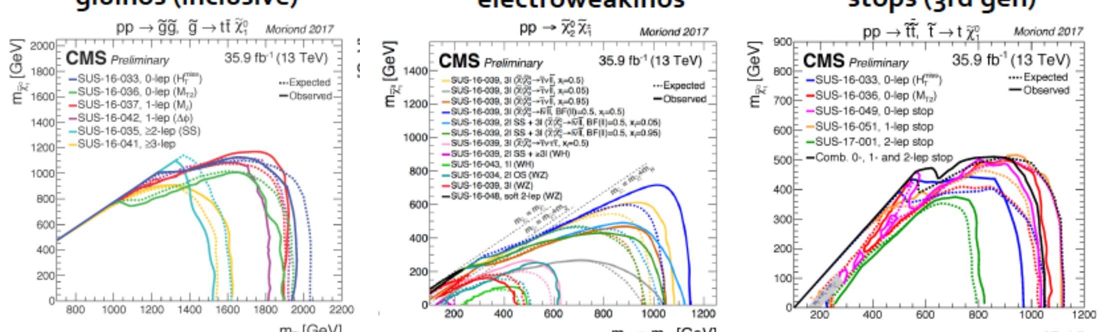
μ(signal strength)= 1.06 +0.25-0.24
Significance 4.9 σ; CMS combined will be > 5 σ

SUSY Searches

Broad program: 19 searches completed with full 2016 CMS dataset, with several already submitted to journals

- Probing different models (inclusive production, strong and electroweak production, and 3rd generation sparticles (stops)
- Different final states(with leptons, photons, jets) and analysis techniques

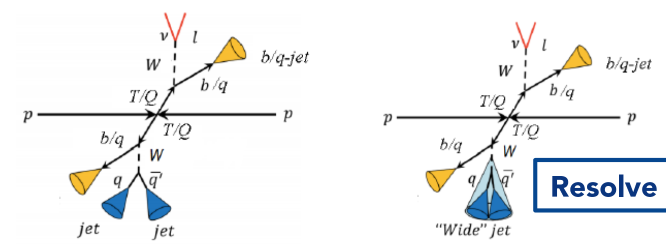
gluinos (inclusive) electroweakinos stops (3rd gen)



Boosted Objects, e.g., Vector Like Quarks

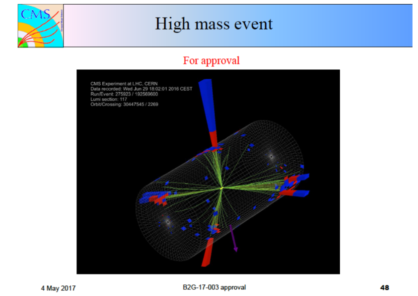
$$T\bar{T} \rightarrow bW^+\bar{b}W^- \rightarrow b\ell\nu\bar{b}q\bar{q}'$$

B2G-17-003

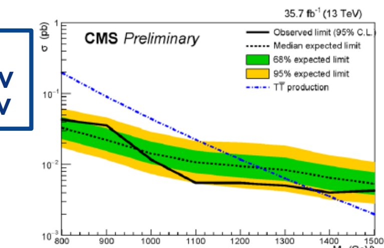


Resolve into two subjets

AK8 jets

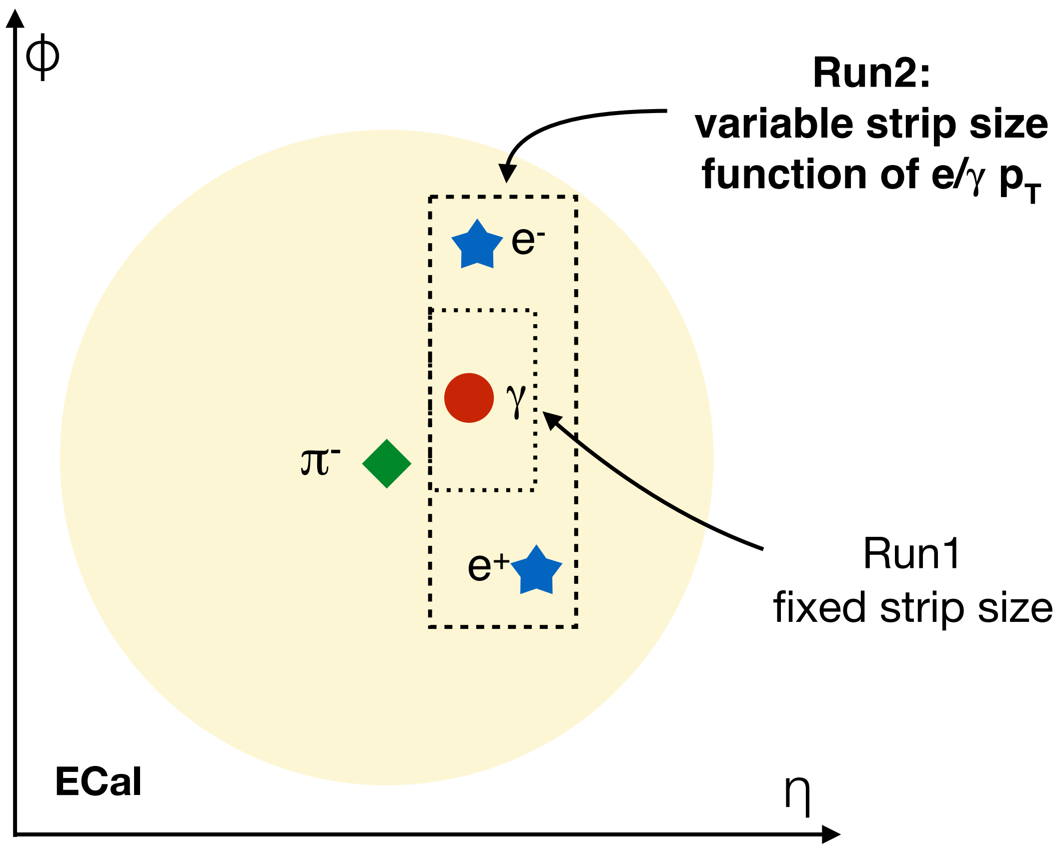


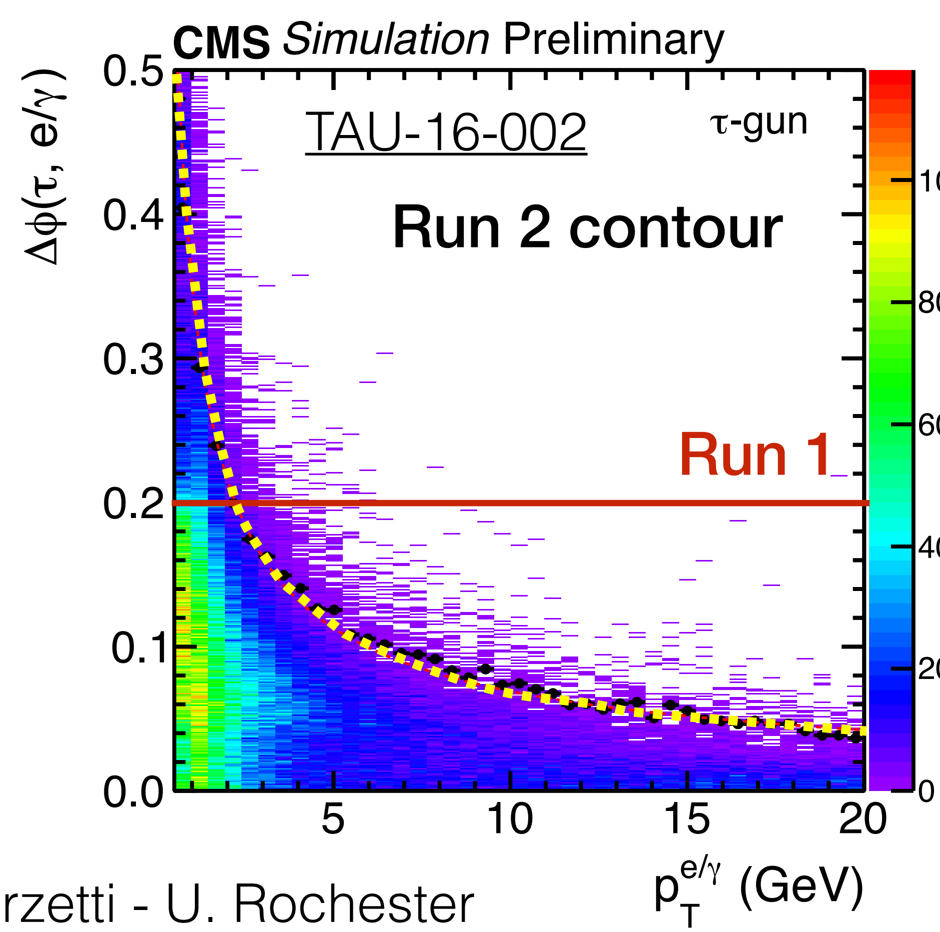
Limit:
Expected: 1245 GeV
Observed: 1365 GeV

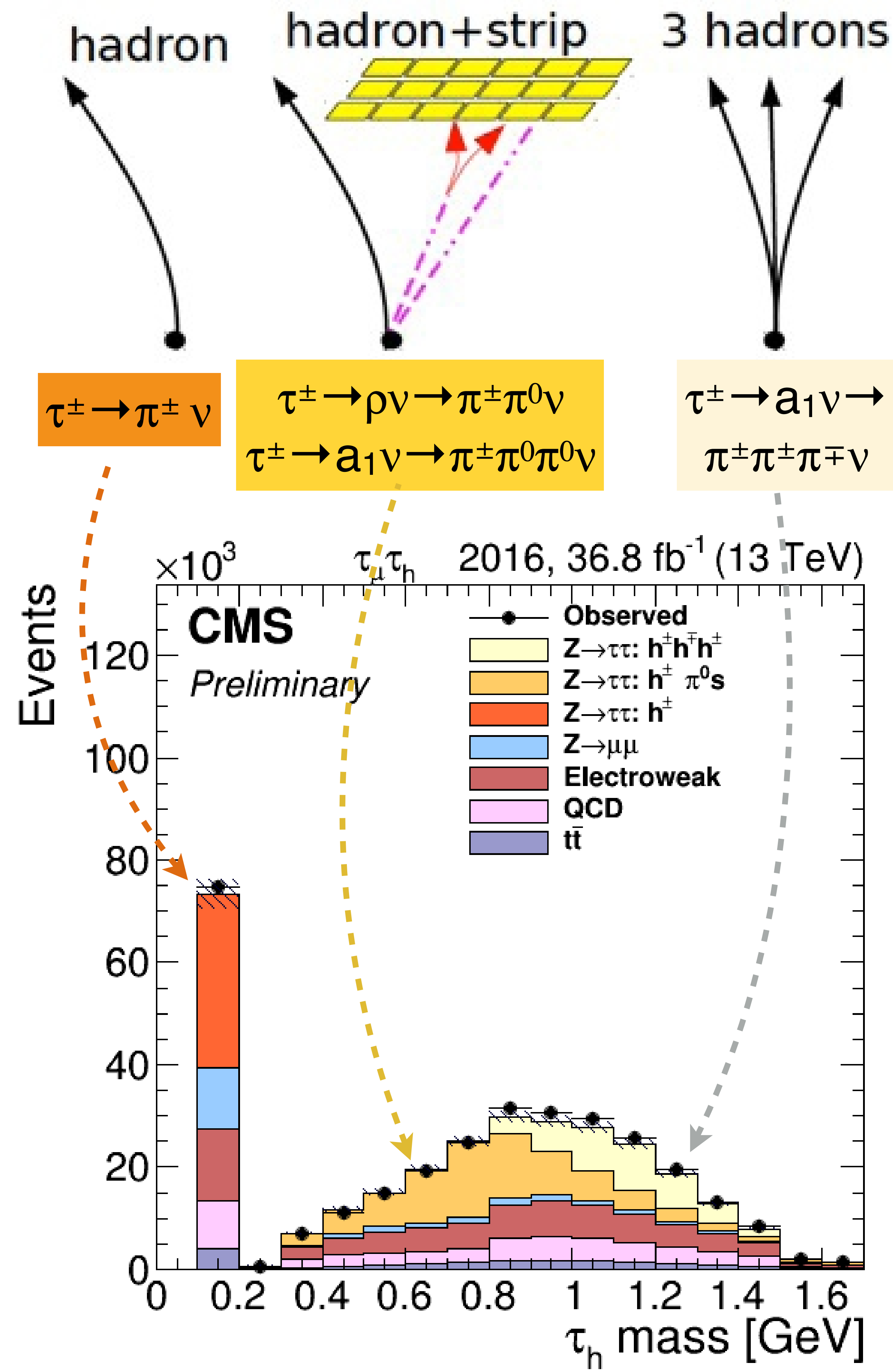


Taus

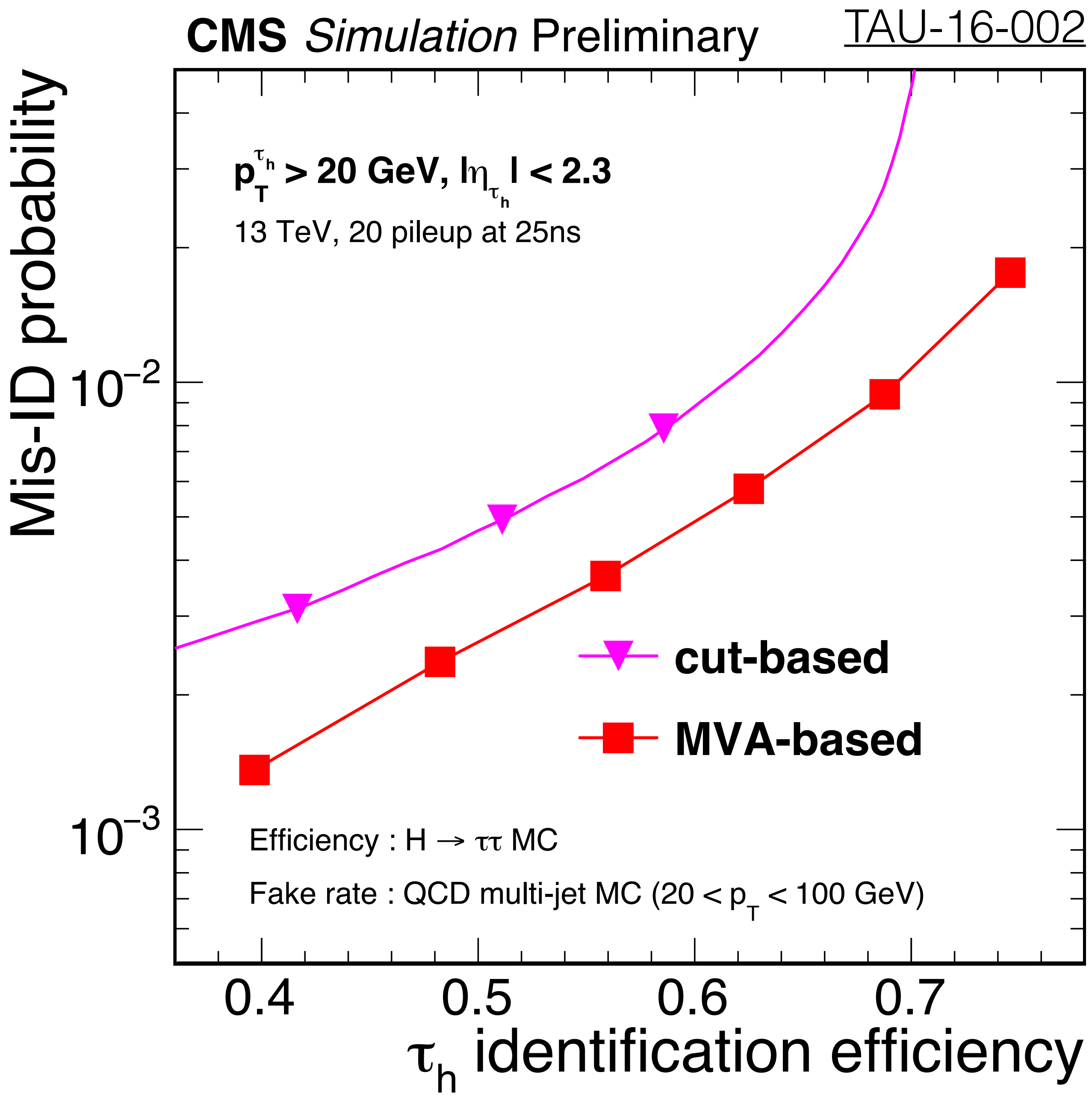
Taus

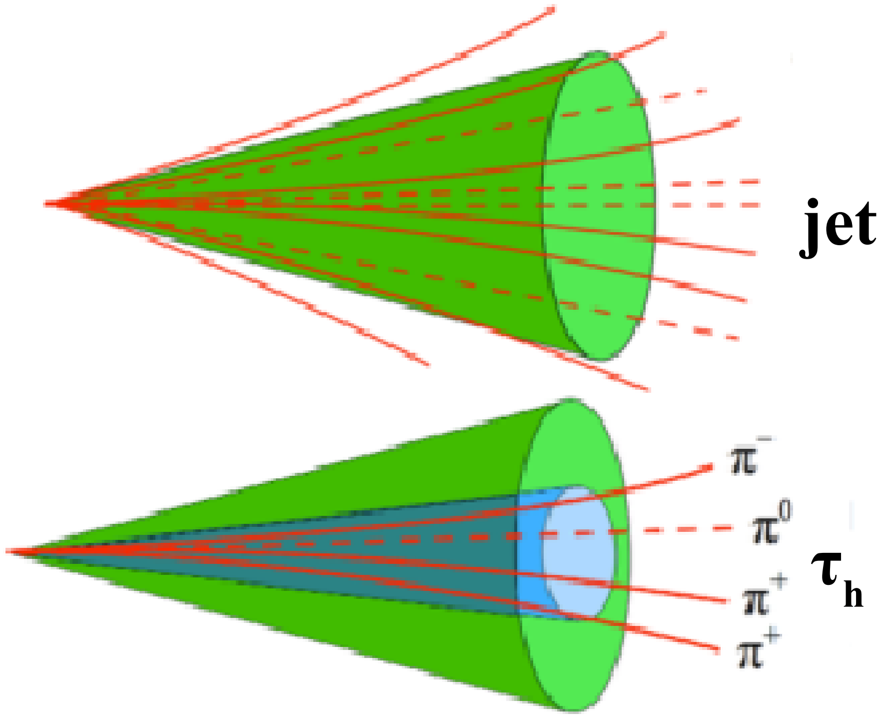

φ
Run2:
variable strip size
function of e/γ p_T
e-
γ
π-
e+
Run1
fixed strip size
ECal
η


CMS Simulation Preliminary
TAU-16-002
τ-gun
Run 2 contour
Run 1
Δφ(τ, e/γ)
$p_T^{e/\gamma}$ (GeV)


hadron
hadron+strip
3 hadrons
$\tau^\pm \to \pi^\pm \nu$
$\tau^\pm \to \rho\nu \to \pi^\pm\pi^0\nu$
$\tau^\pm \to a_1\nu \to \pi^\pm\pi^0\pi^0\nu$
$\tau^\pm \to a_1\nu \to \pi^\pm\pi^\pm\pi^\mp\nu$
$\times 10^3$
$\tau_\mu\tau_h$ 2016, 36.8 fb^{-1} (13 TeV)
CMS
Preliminary
Observed
$Z\to\tau\tau$: $h^\pm h^\mp h^\pm$
$Z\to\tau\tau$: $h^\pm$ π^0s
$Z\to\tau\tau$: $h^\pm$
$Z\to\mu\mu$
Electroweak
QCD
$t\bar{t}$
Events
τ_h mass [GeV]

Taus





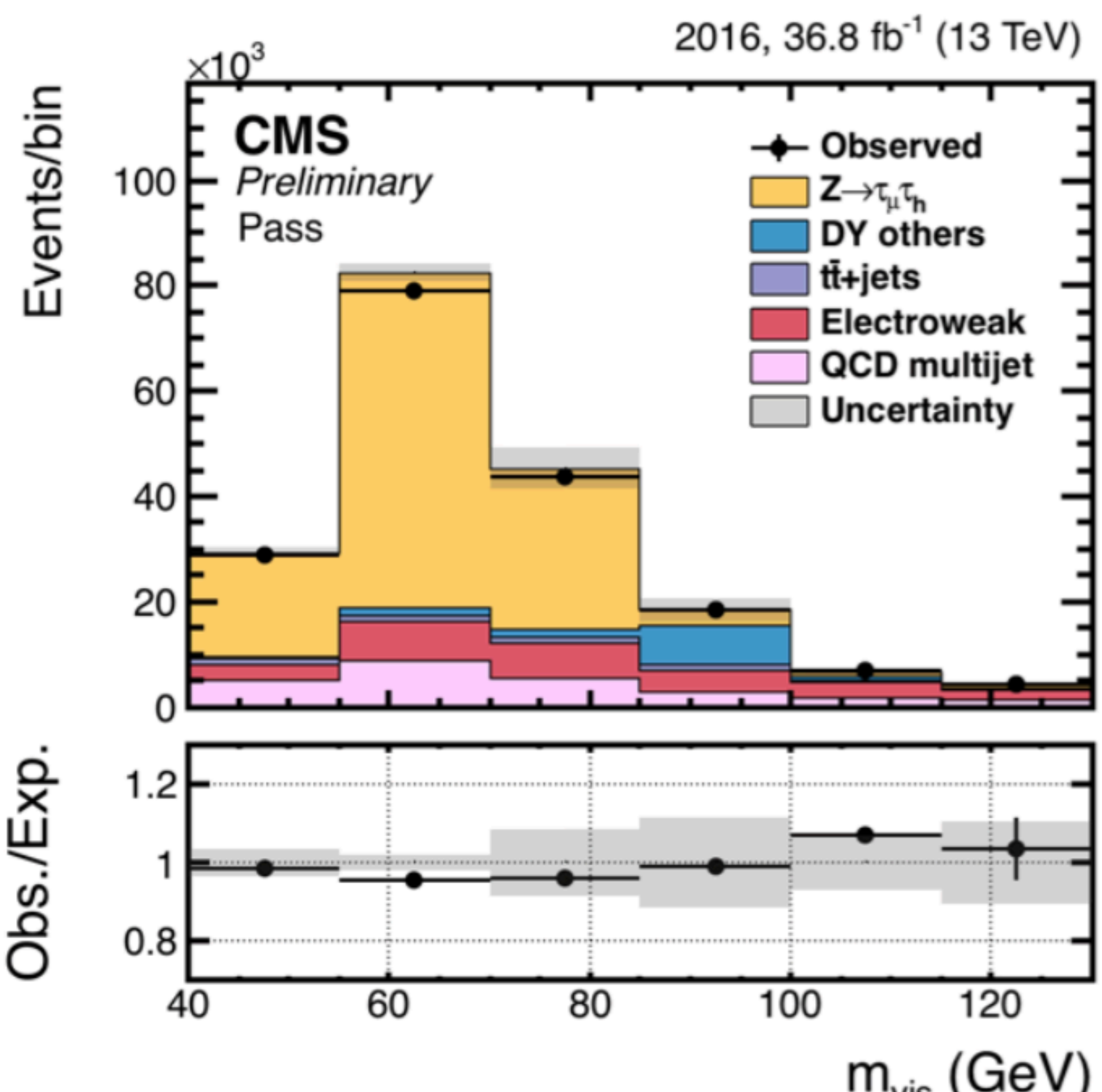
Traditional cut-based approach on isolation.

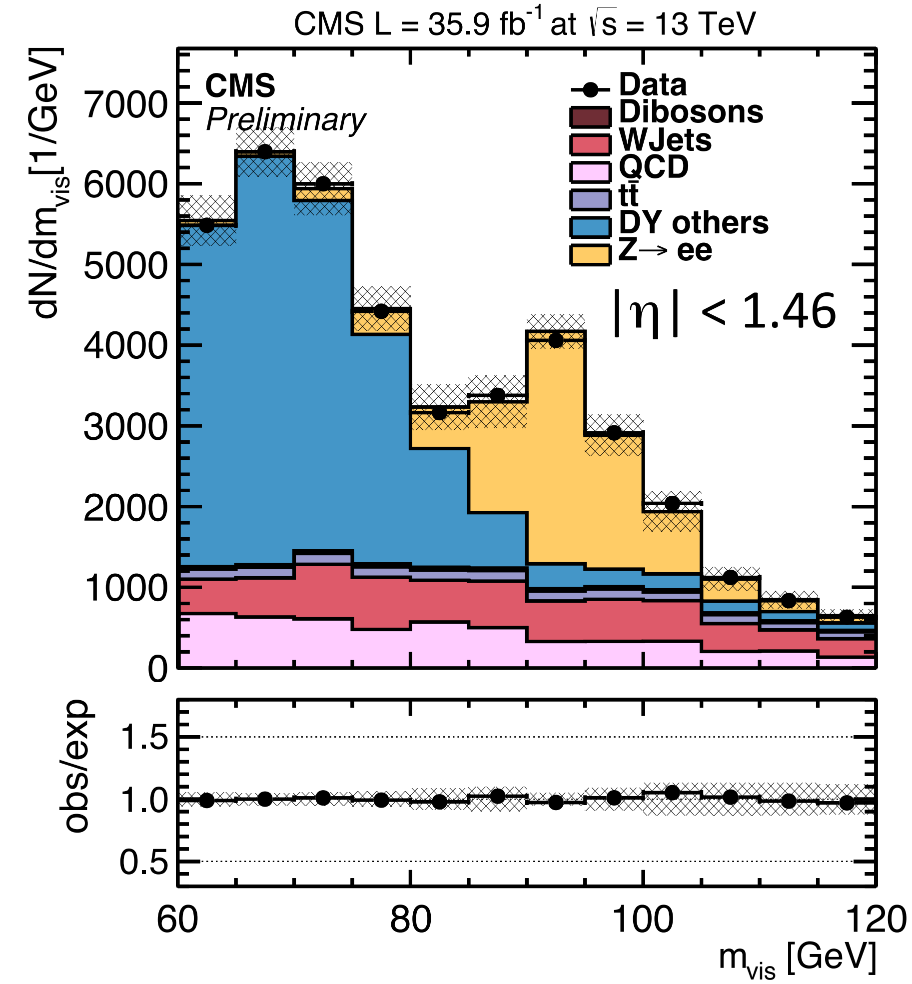
Additional MVA (BDT) approach including strip shape and lifetime information

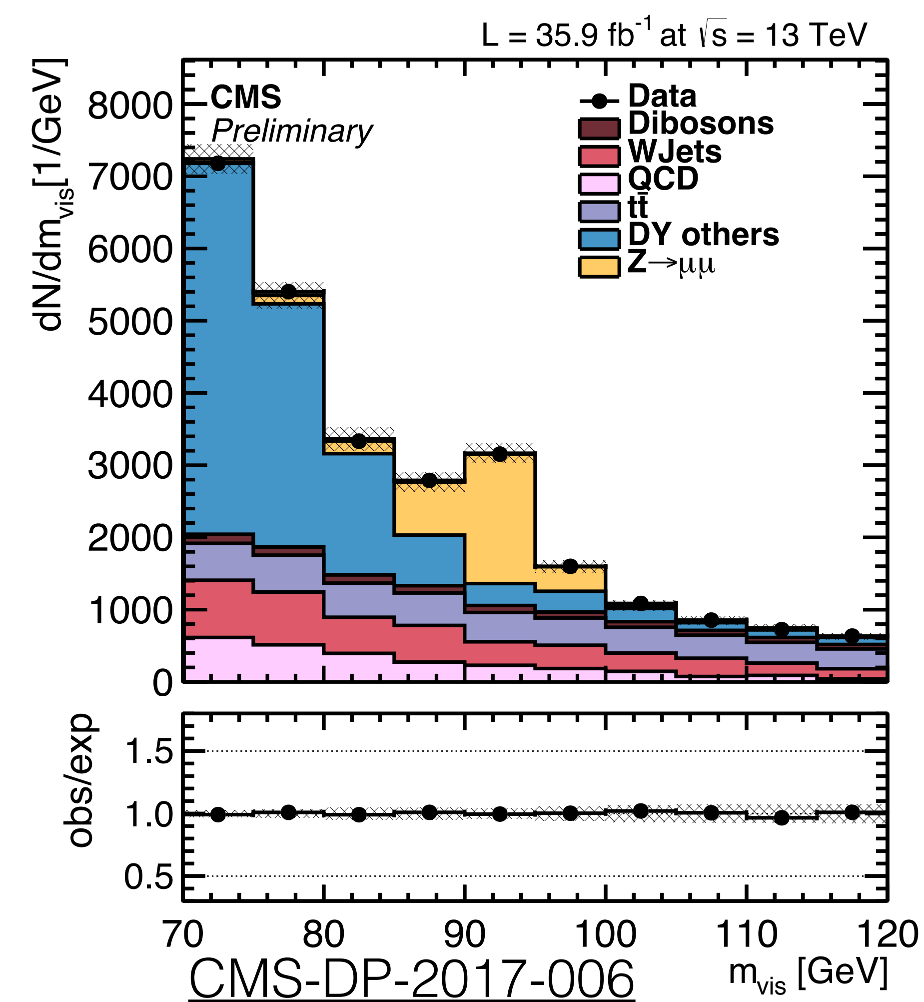
MVA approach outperforms the cut-based one

Taus

Use Z → ττ, ee, μμ as standard candle to measure efficiency and mis-tagging rate



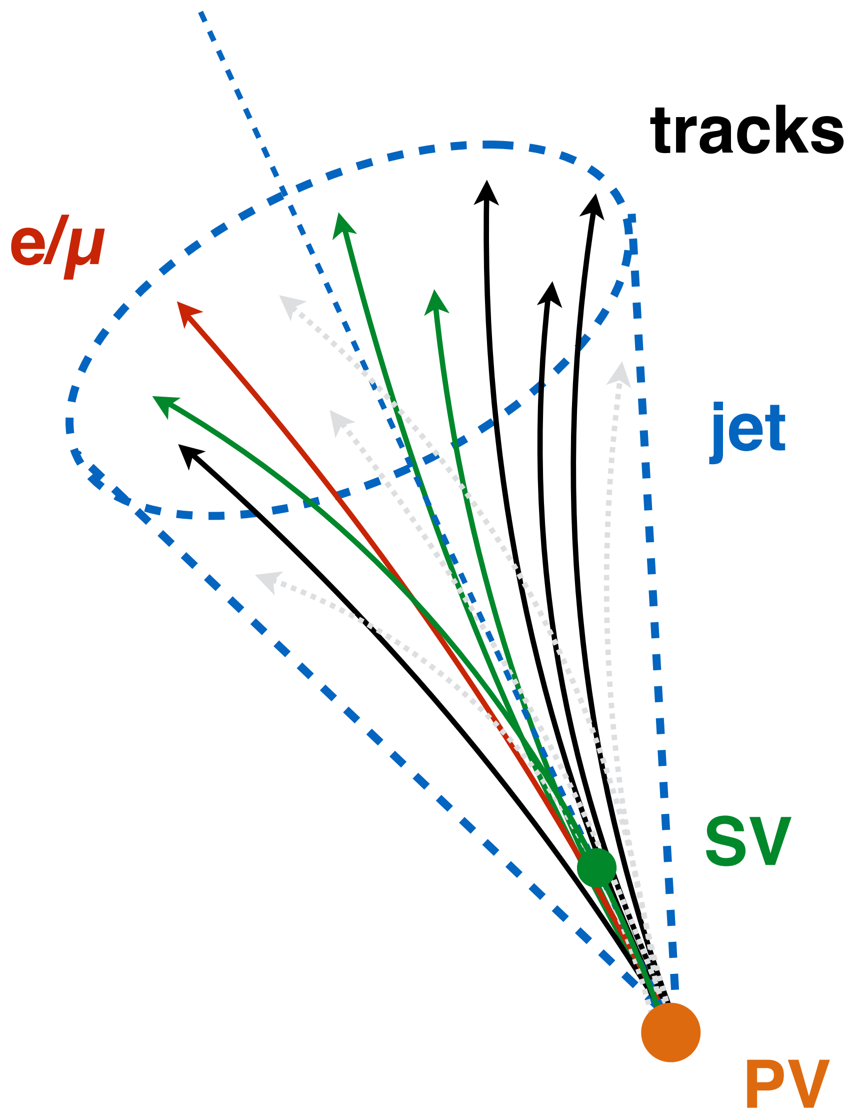


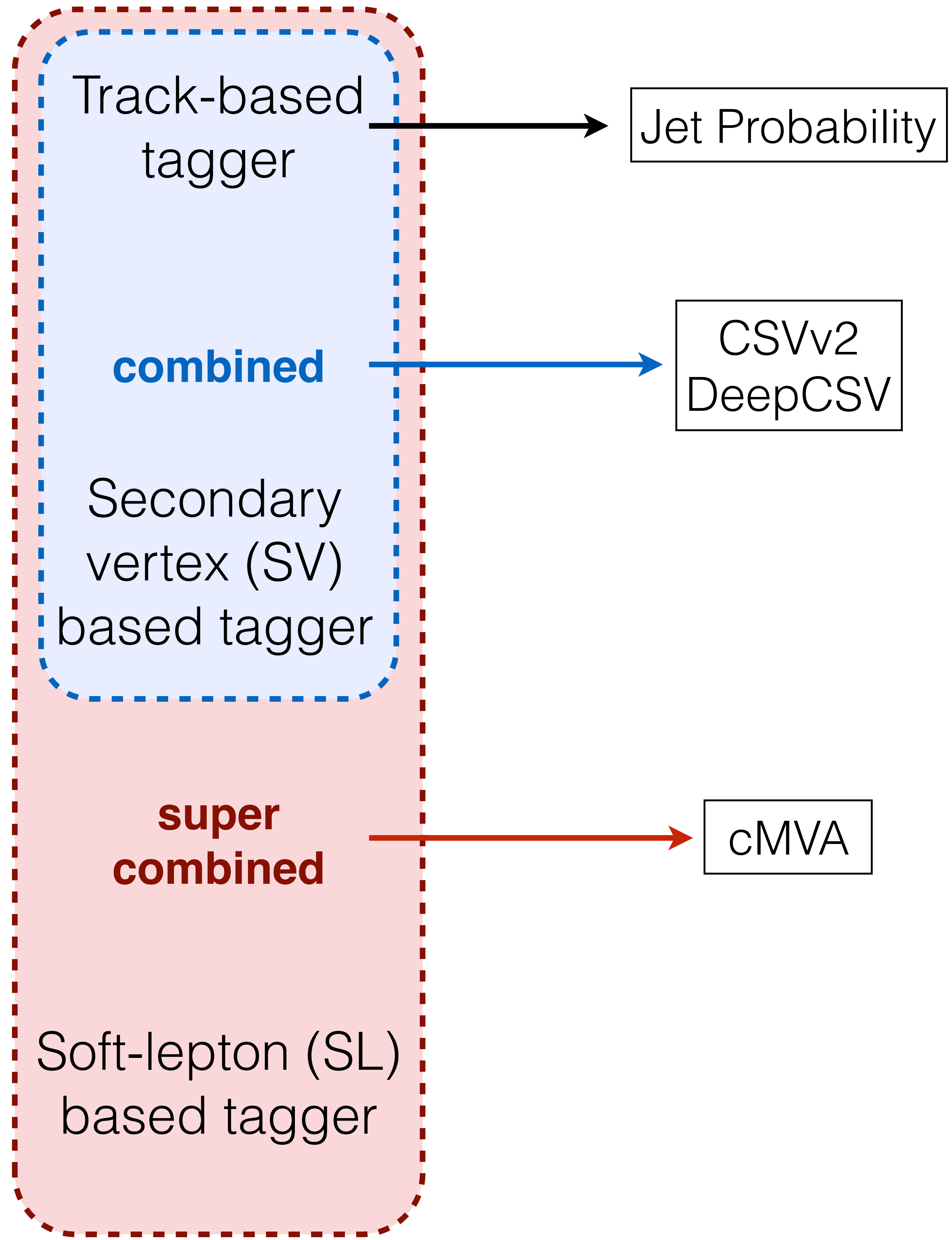


CMS-DP-2017-006

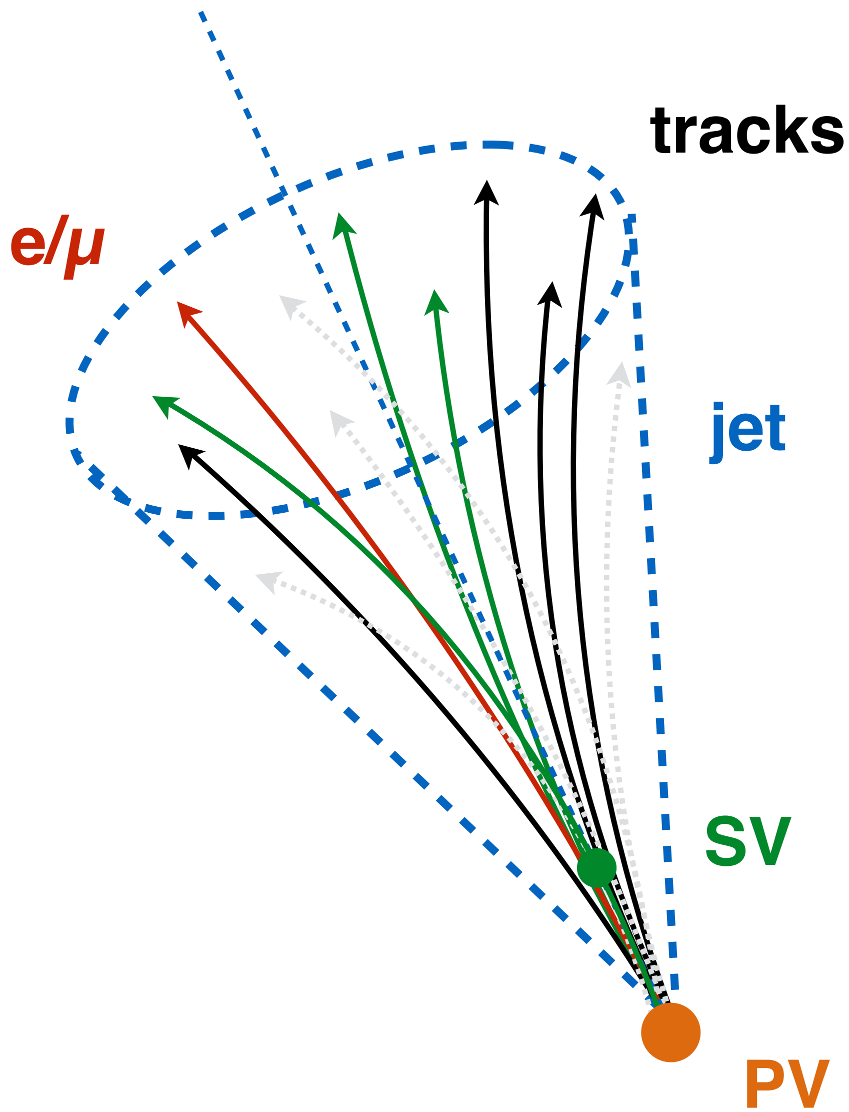
Heavy flavors

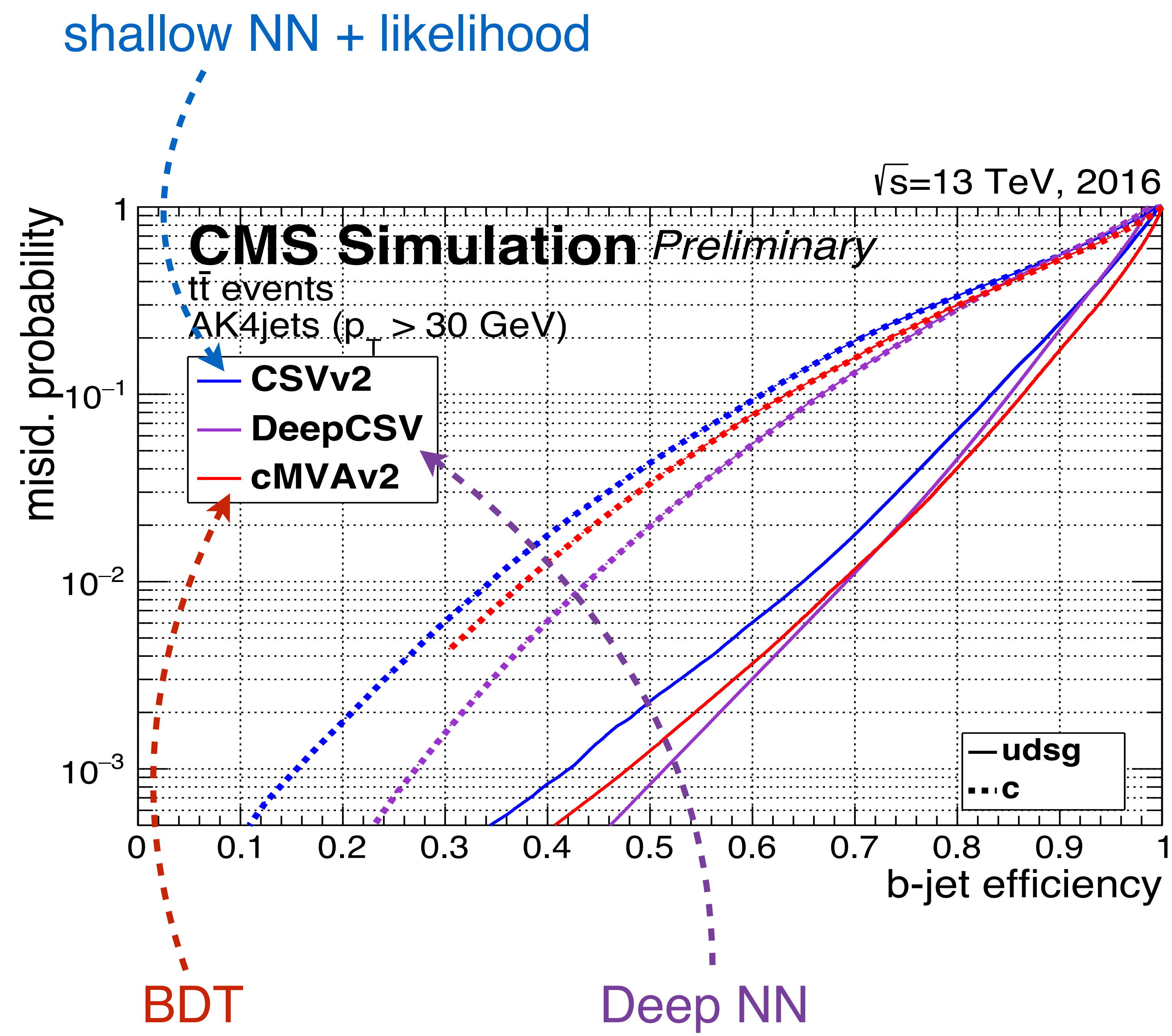
Heavy flavor jets





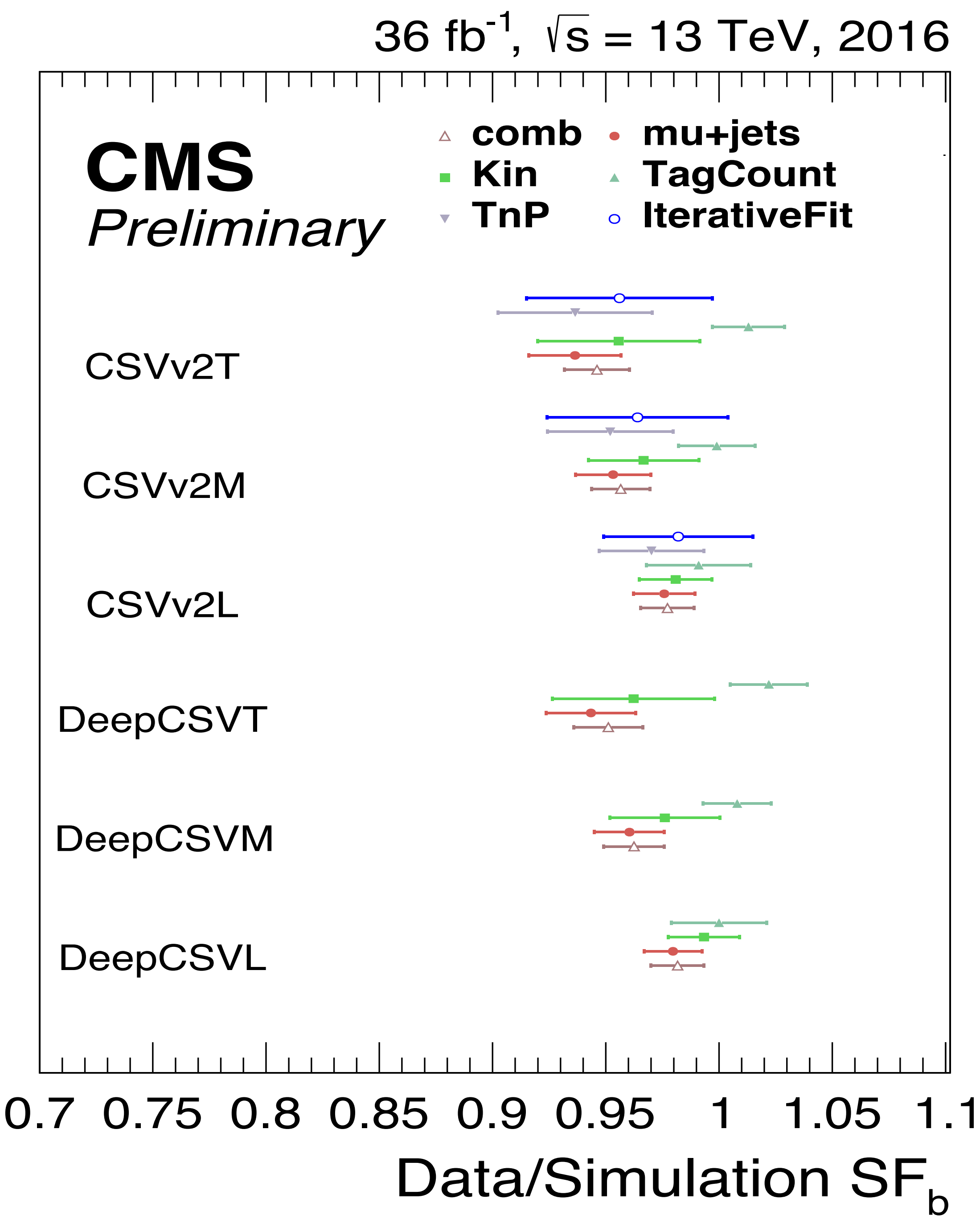
Heavy flavor jets

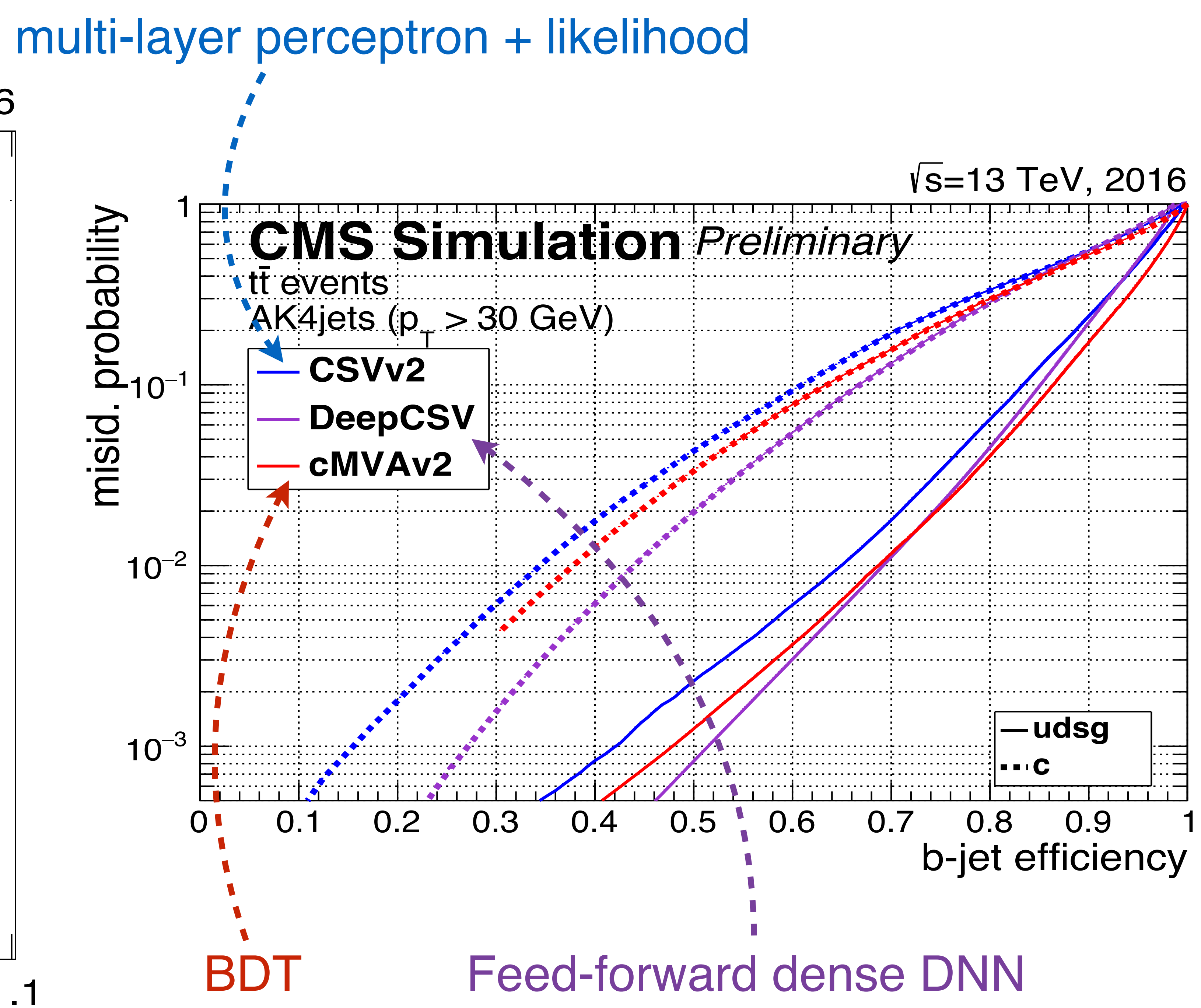




DeepNN shows best performance

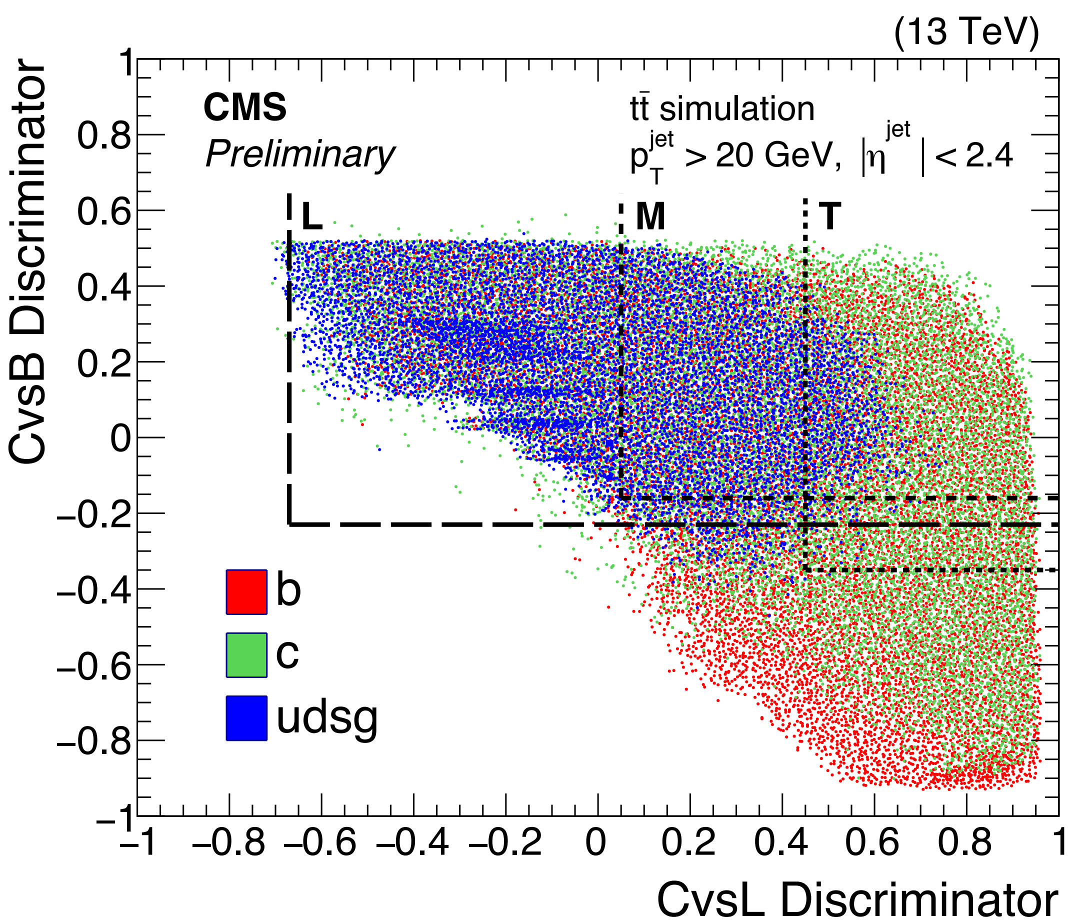
Heavy flavor jets

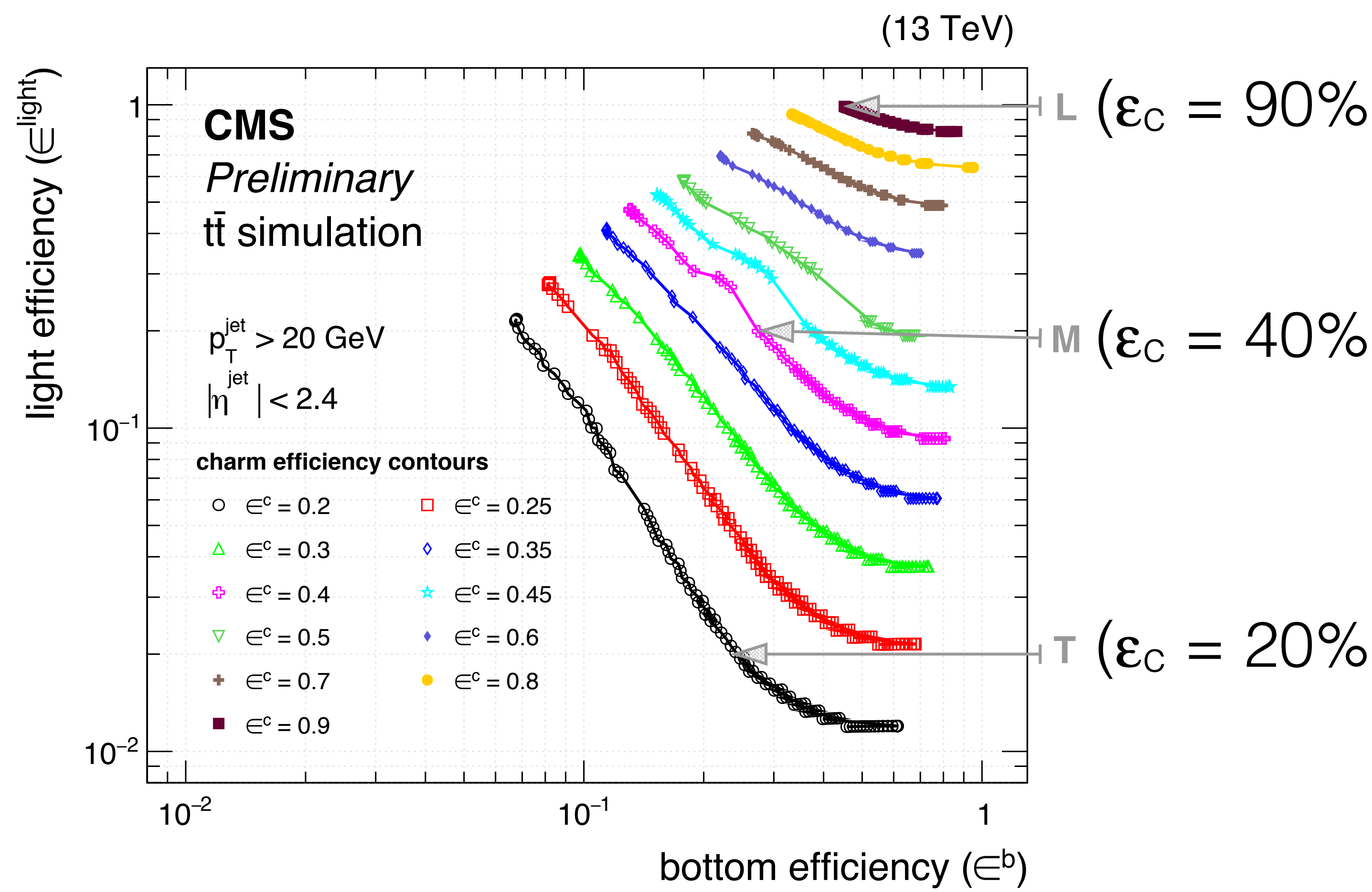




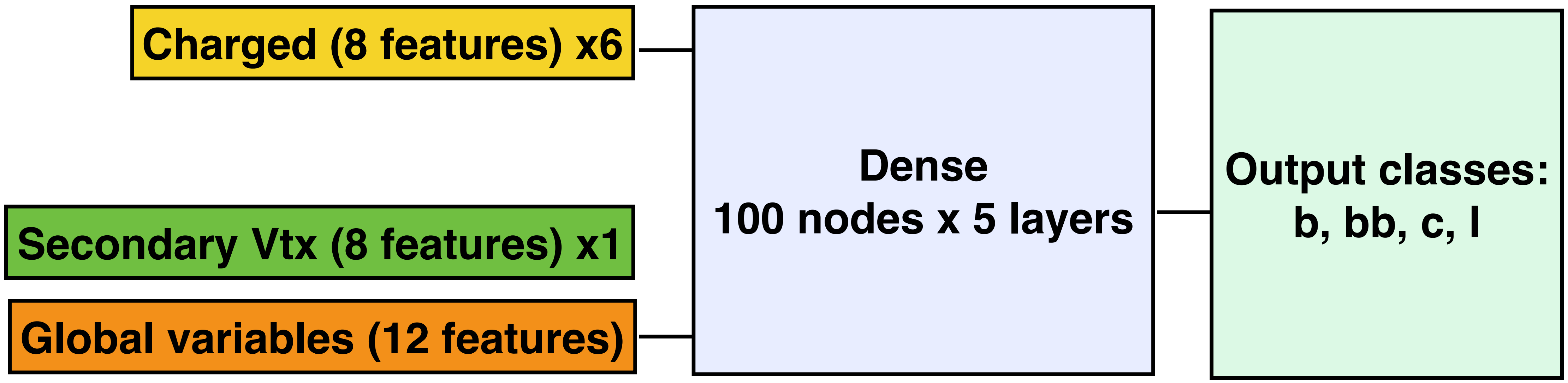
Charm tagging

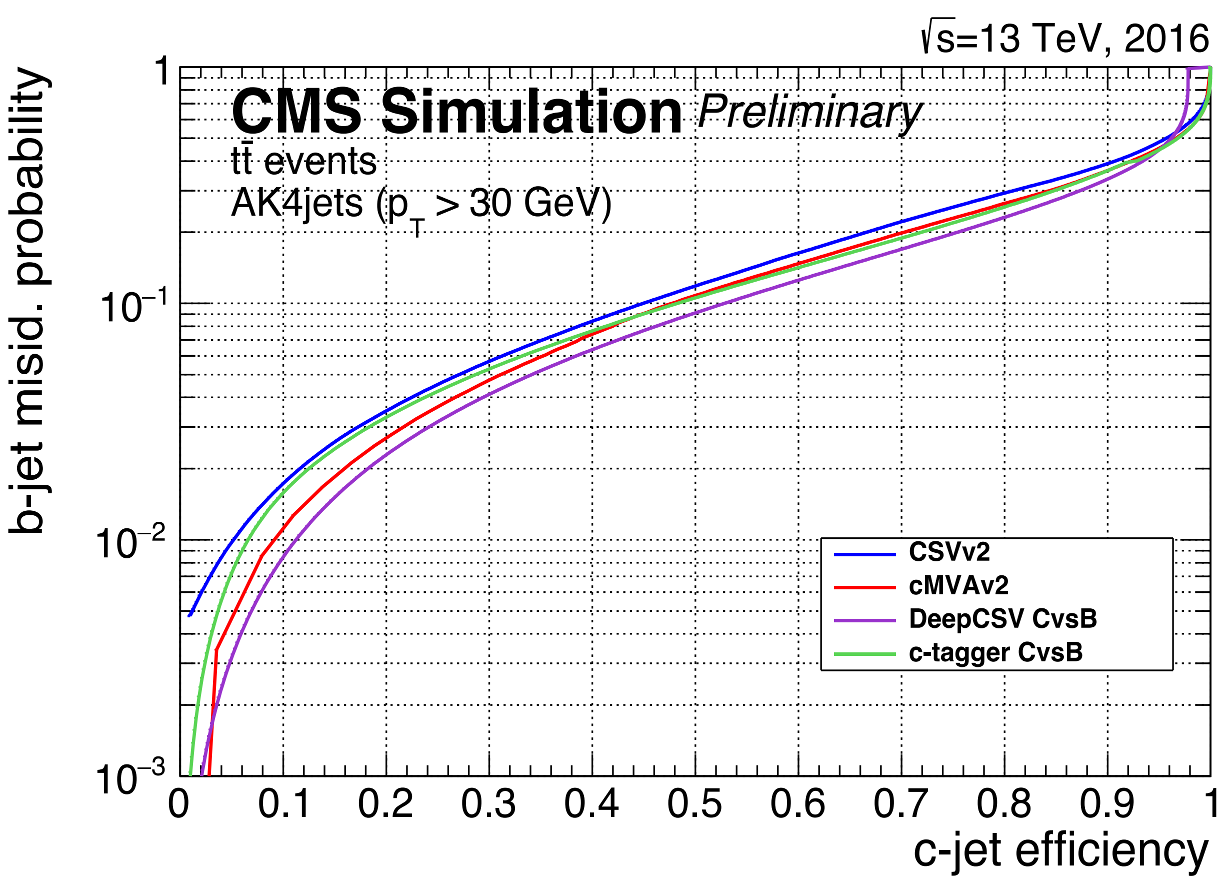
Use two BDT's to discriminate the charm from the light and B components

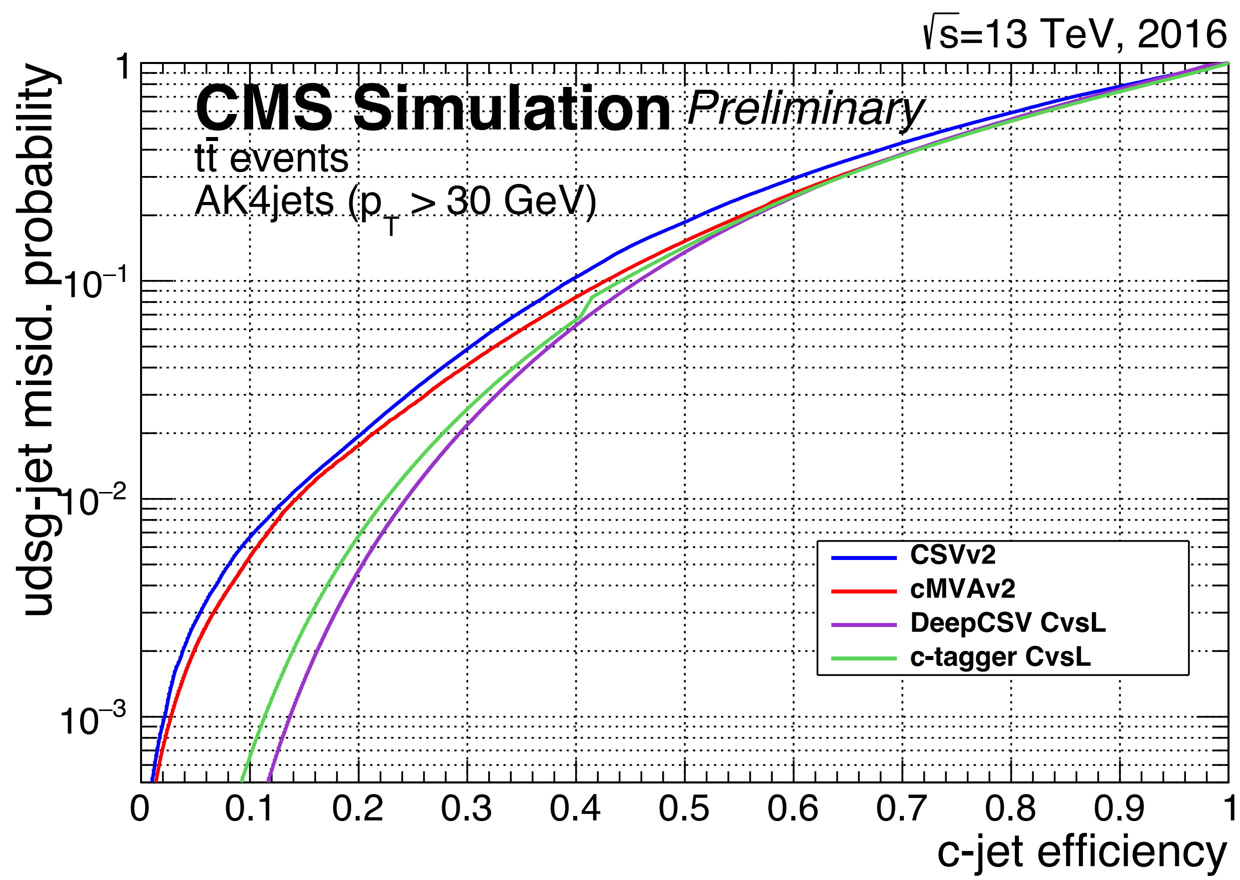




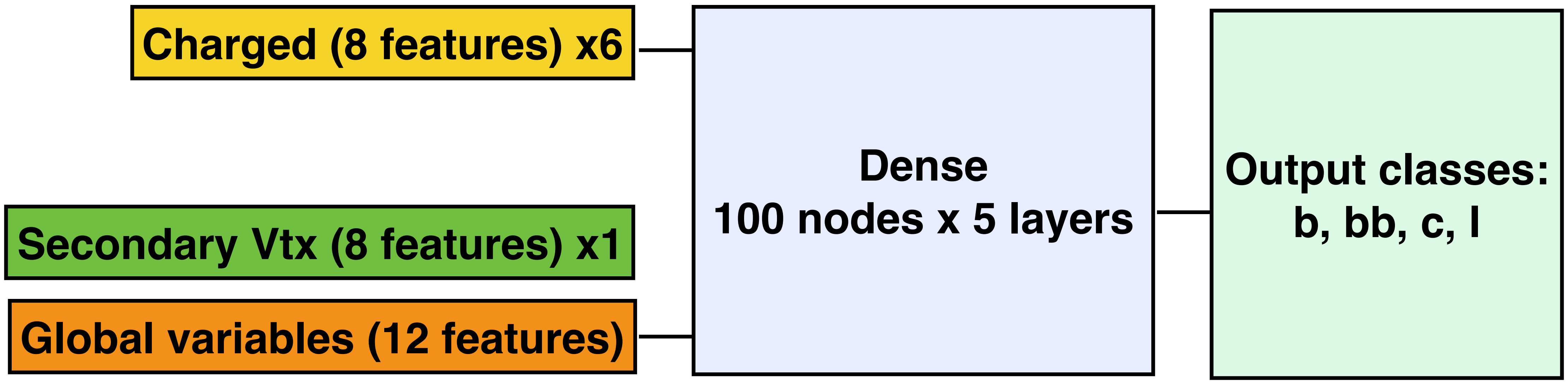
Heavy flavor jets with DNN — DeepCSV



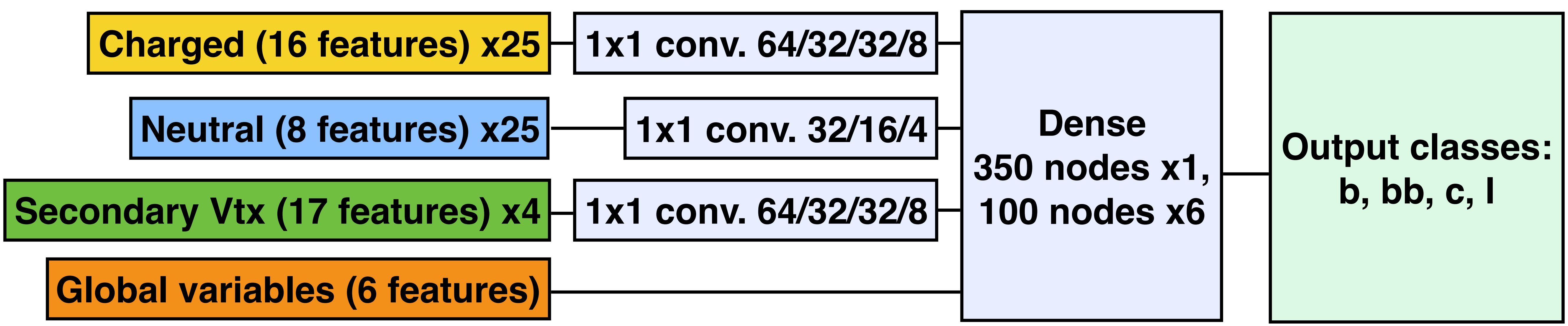




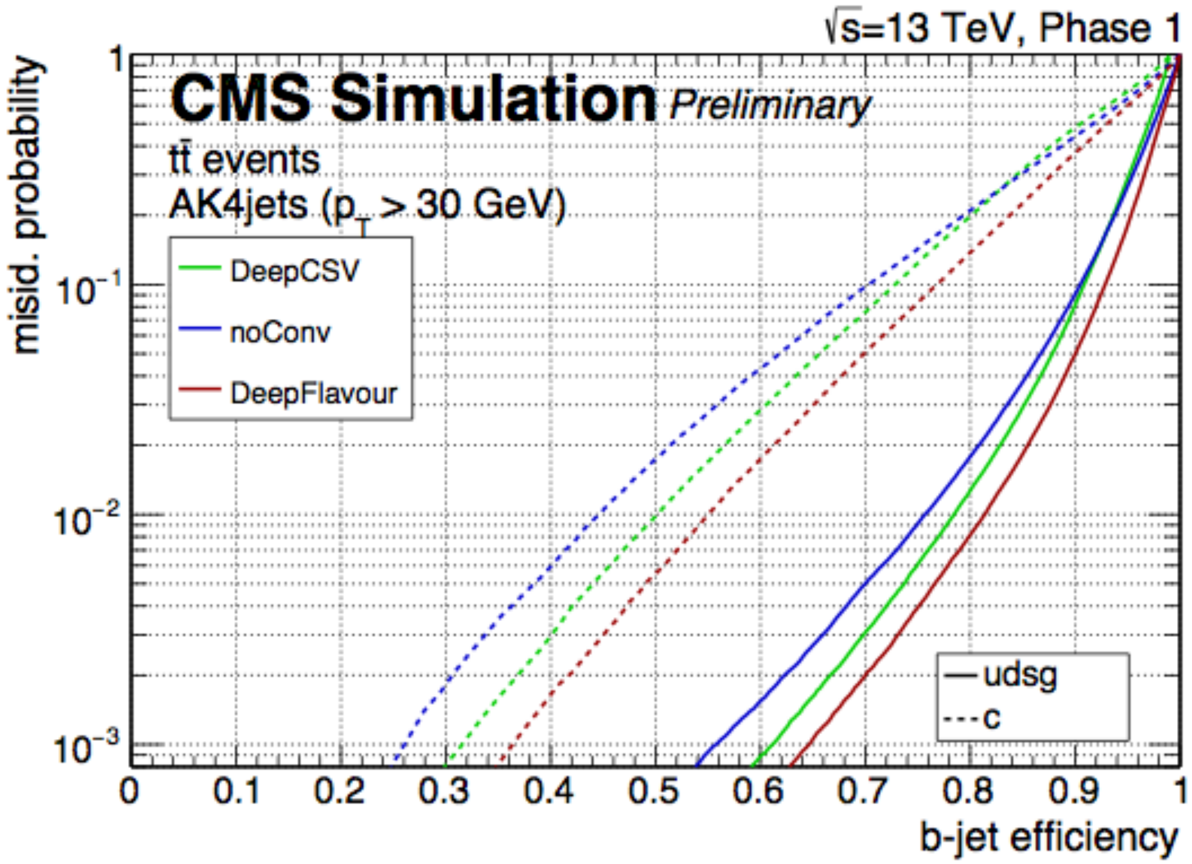
Heavy flavor jets with DNN — Trying deeper networks

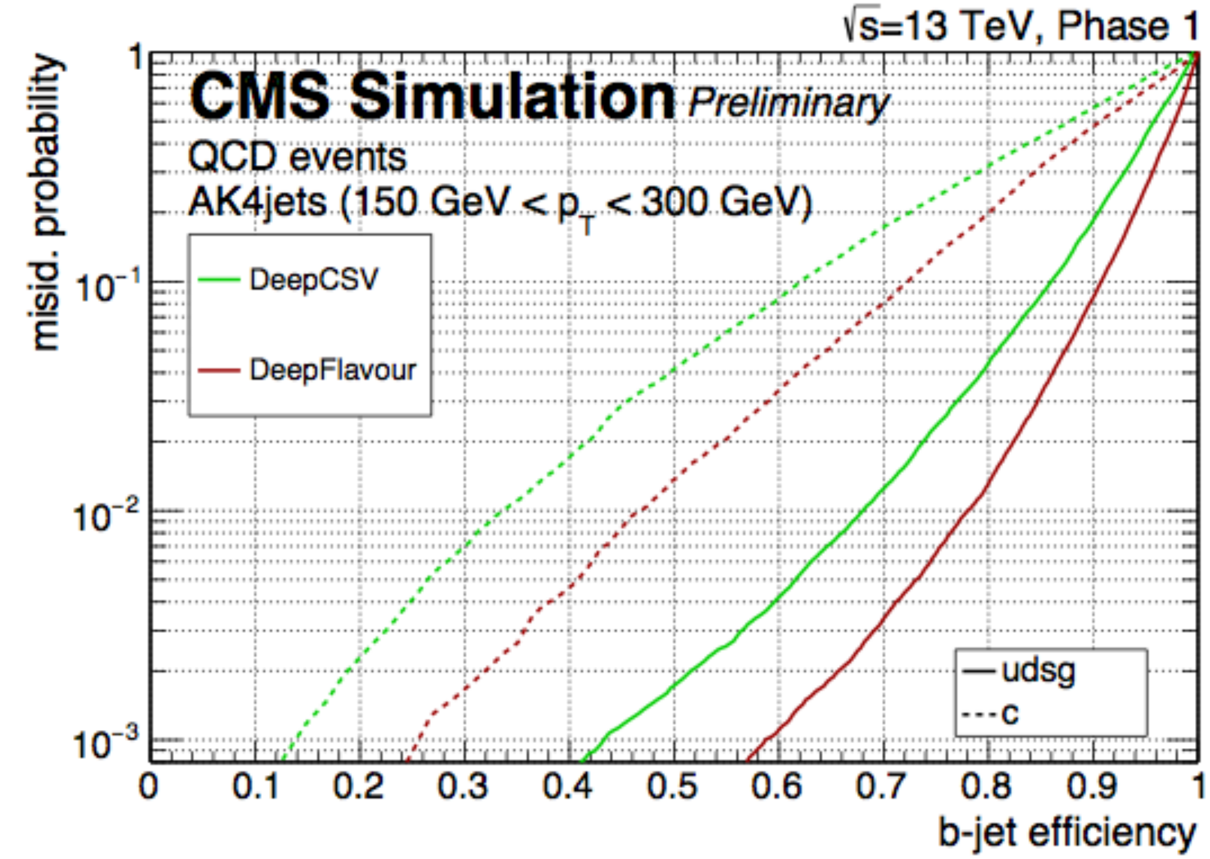


Deep Flavor

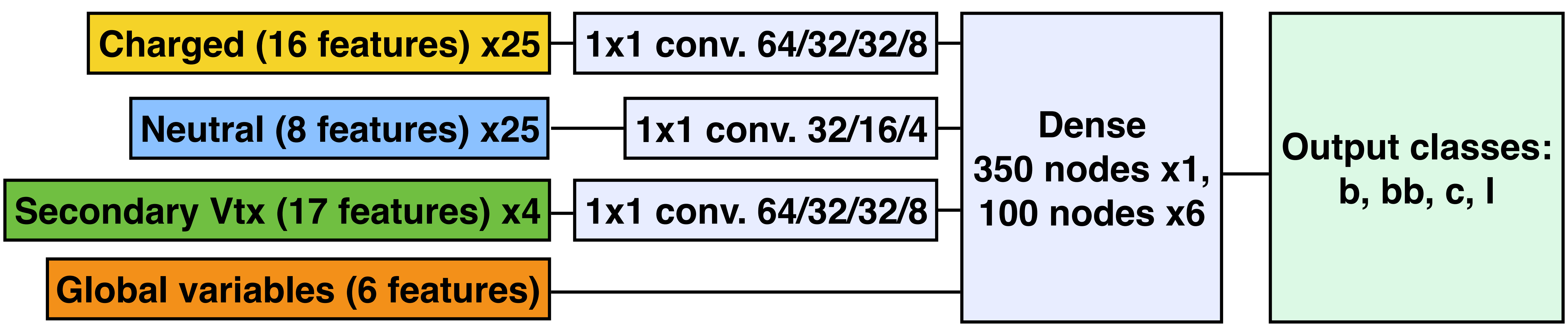


Heavy flavor jets with DNN — Trying deeper networks





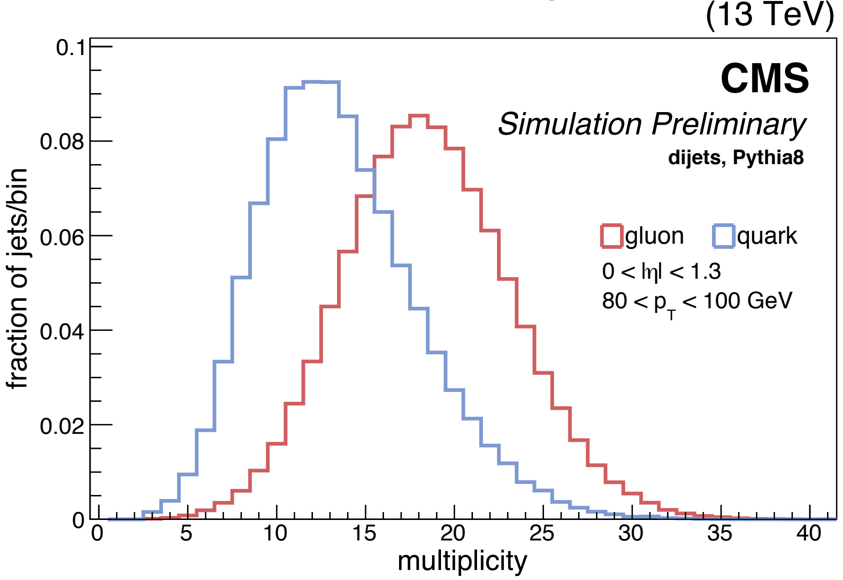
Deep Flavor



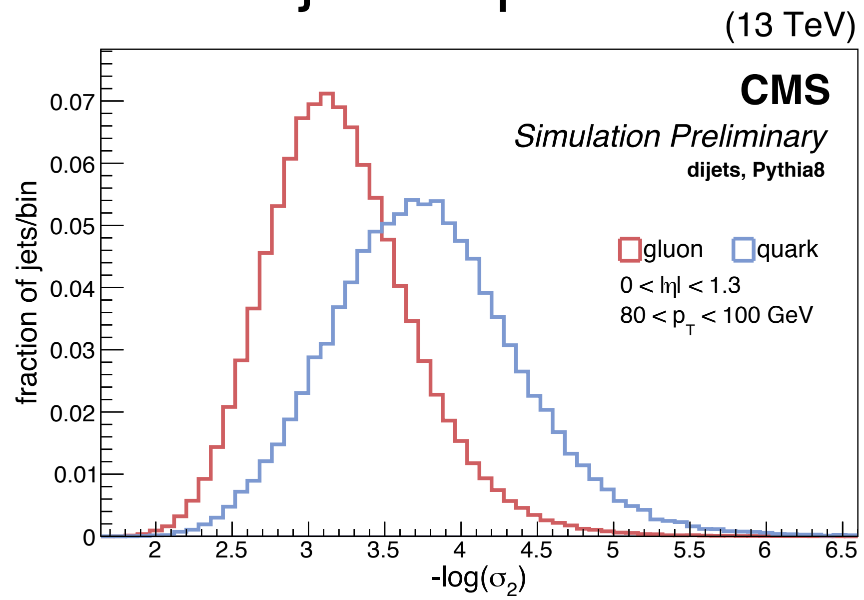
Quarks vs. gluons

Quark / gluon likelihood

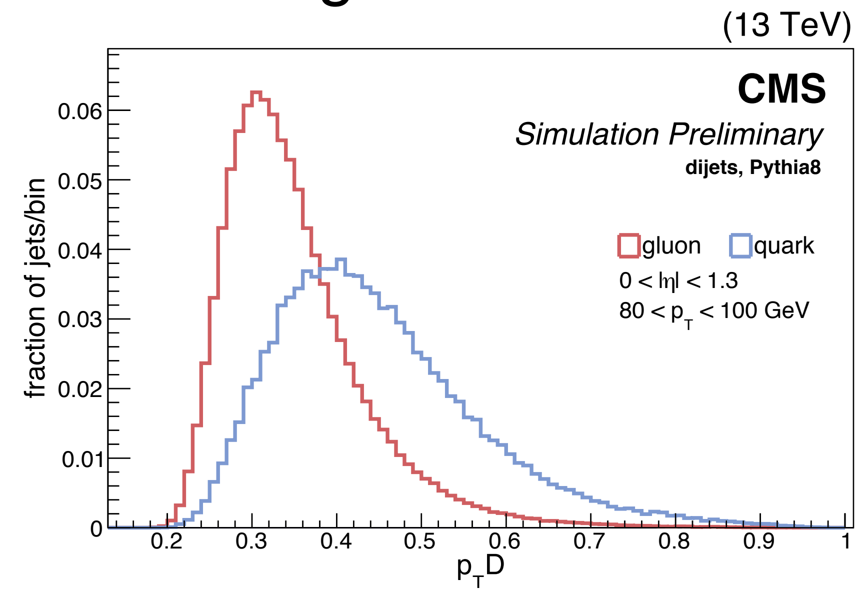
multiplicity

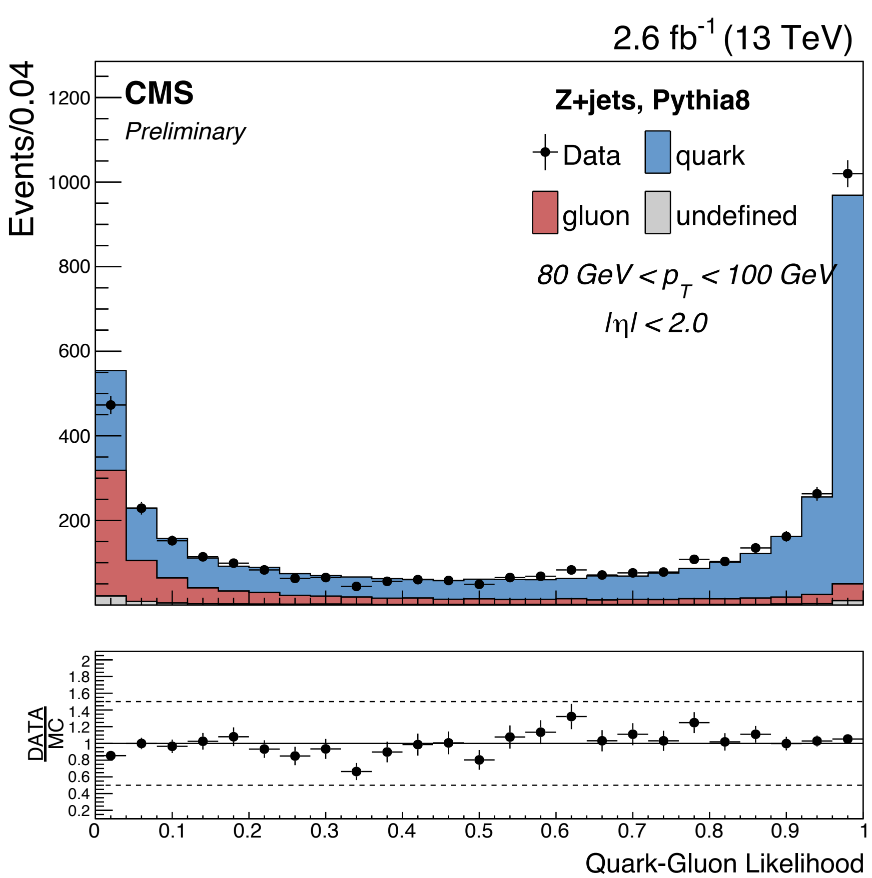


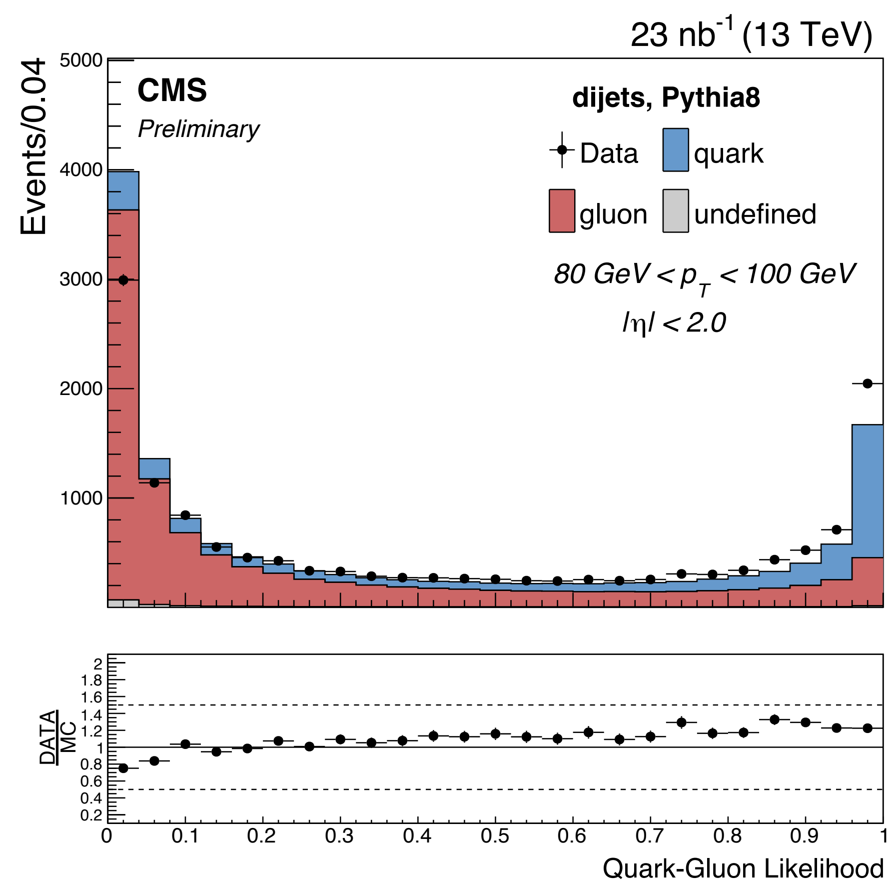
jet shape



fragmentation

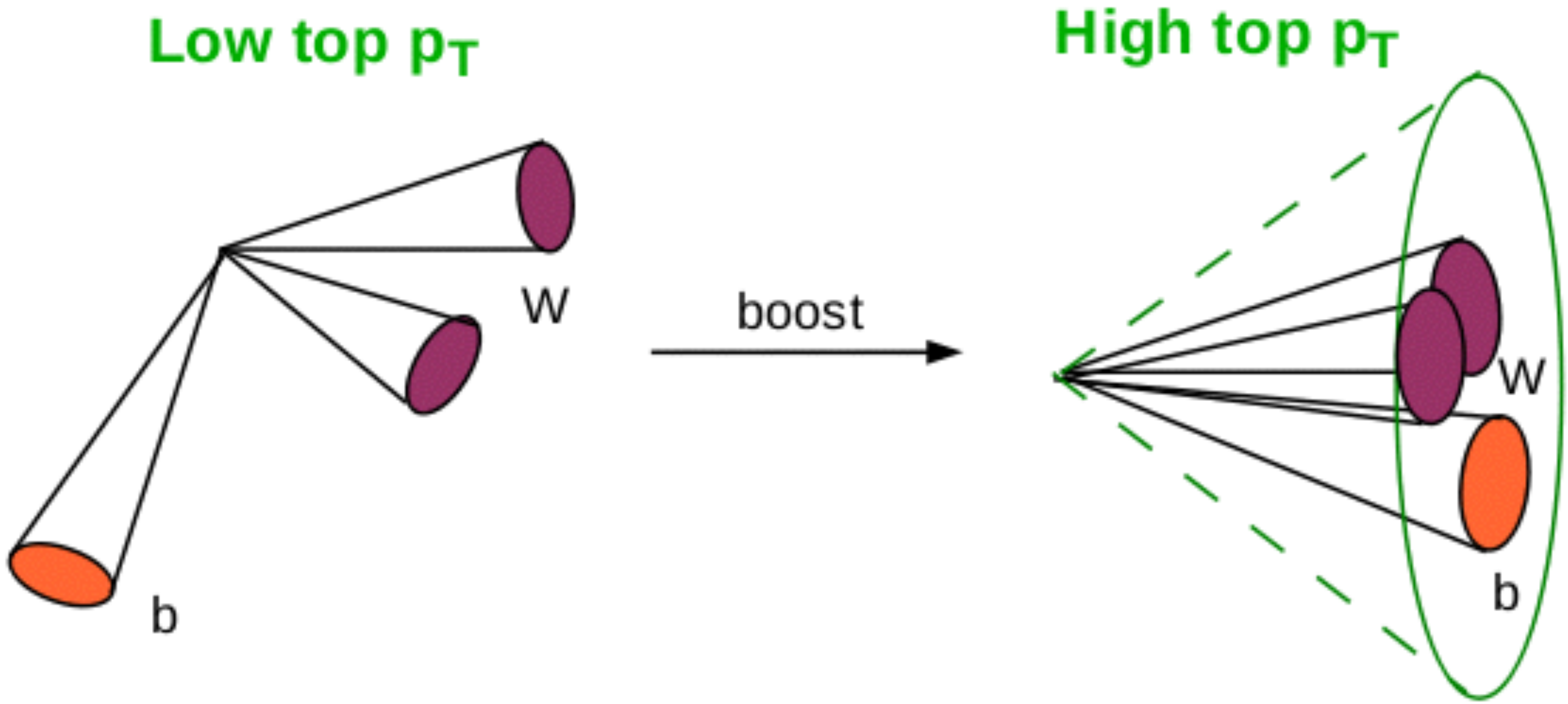






Boosted resonances

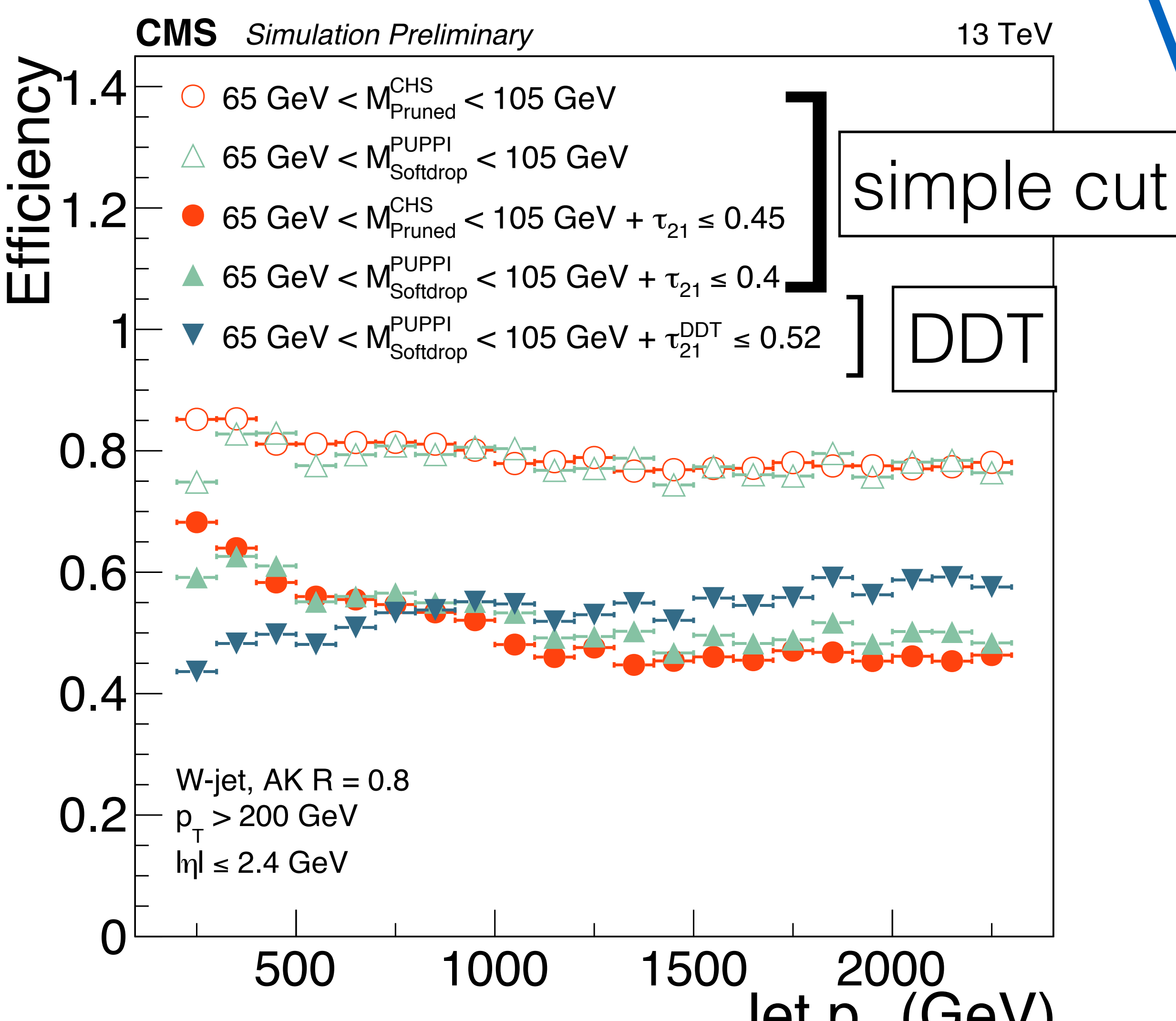
Boosted objects

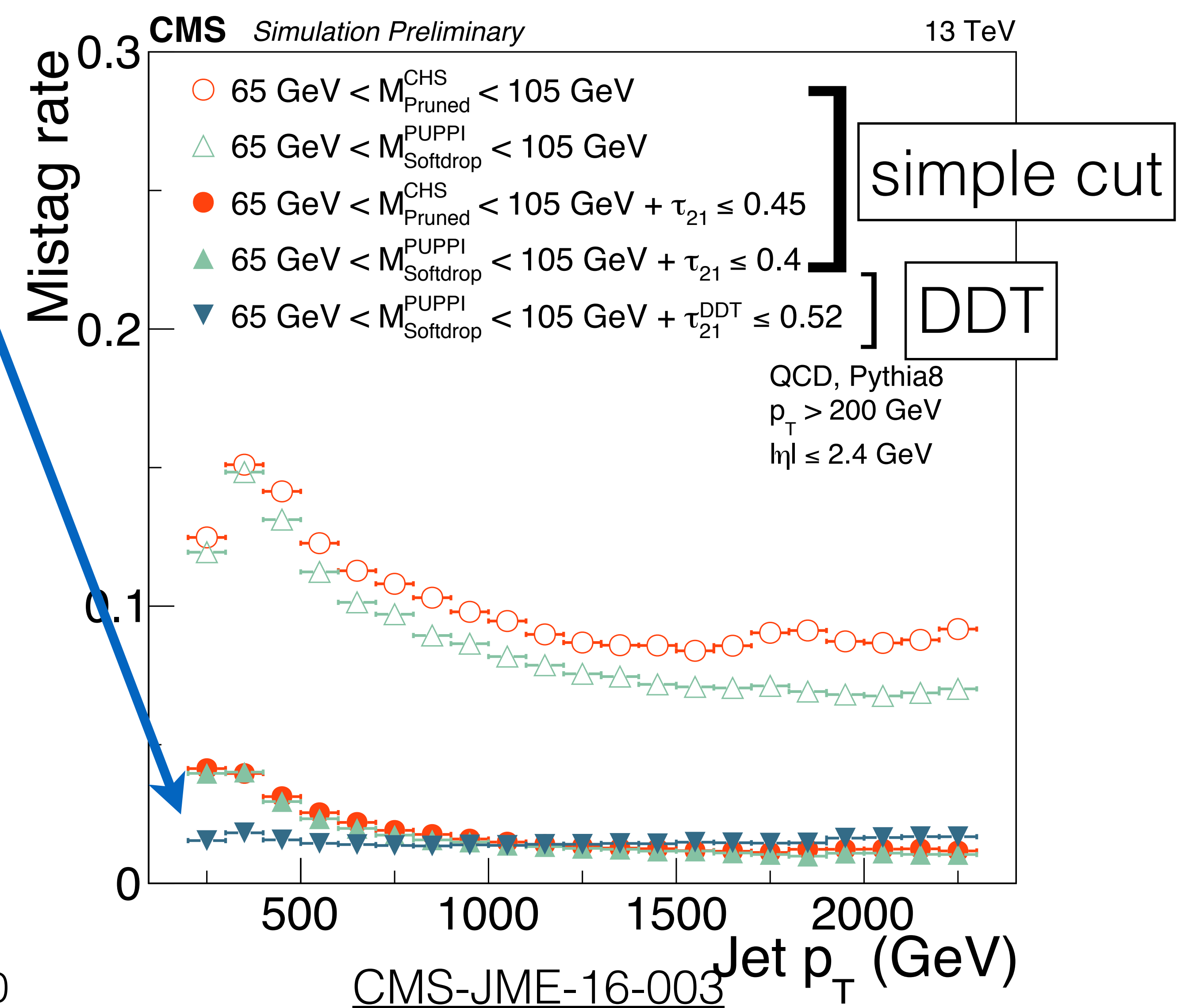


Boosted objects - W/Z tagging

Cut-based approach SoftDrop mass and τ_2/τ_1

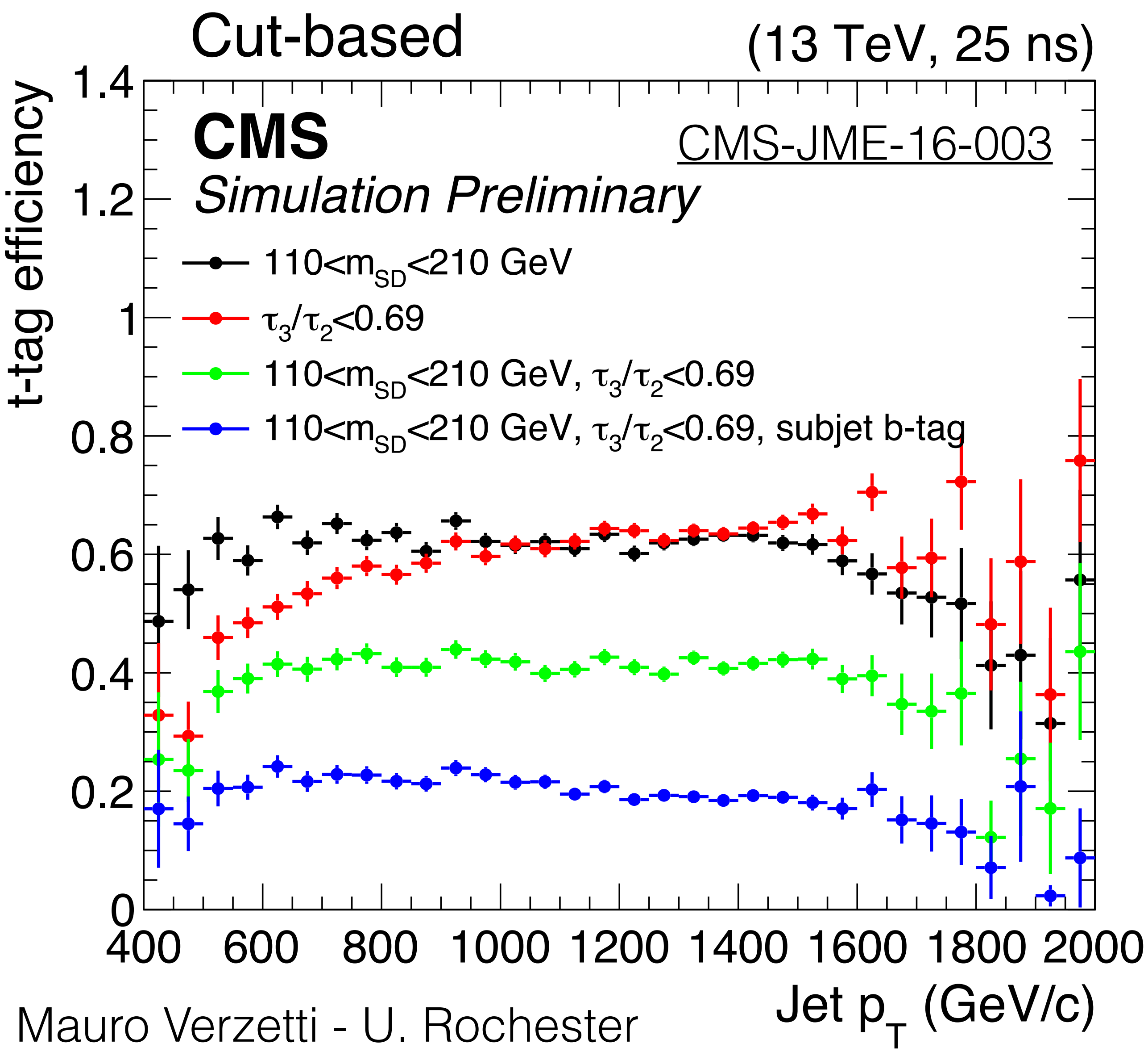
Alternative tagger with mass/p_T decorrelated (DDT approach)
Flat fake rate (advantage for background estimates)

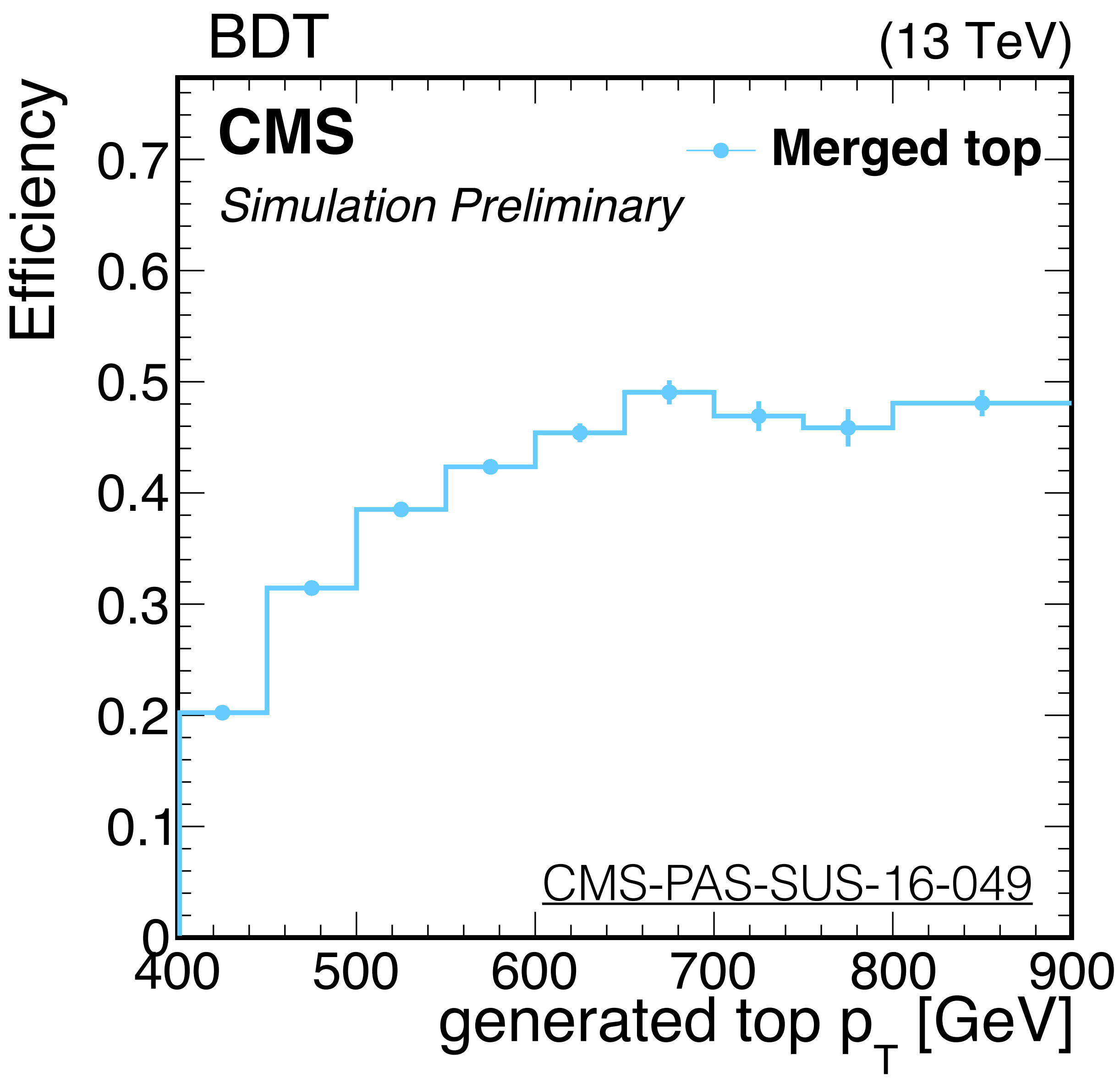




Boosted objects - top

Multiple approaches (cut-based or MVA) leveraging SoftDrop mass, τ_3/τ_2, sub-jet pair masses, and sub-jet b-tagging



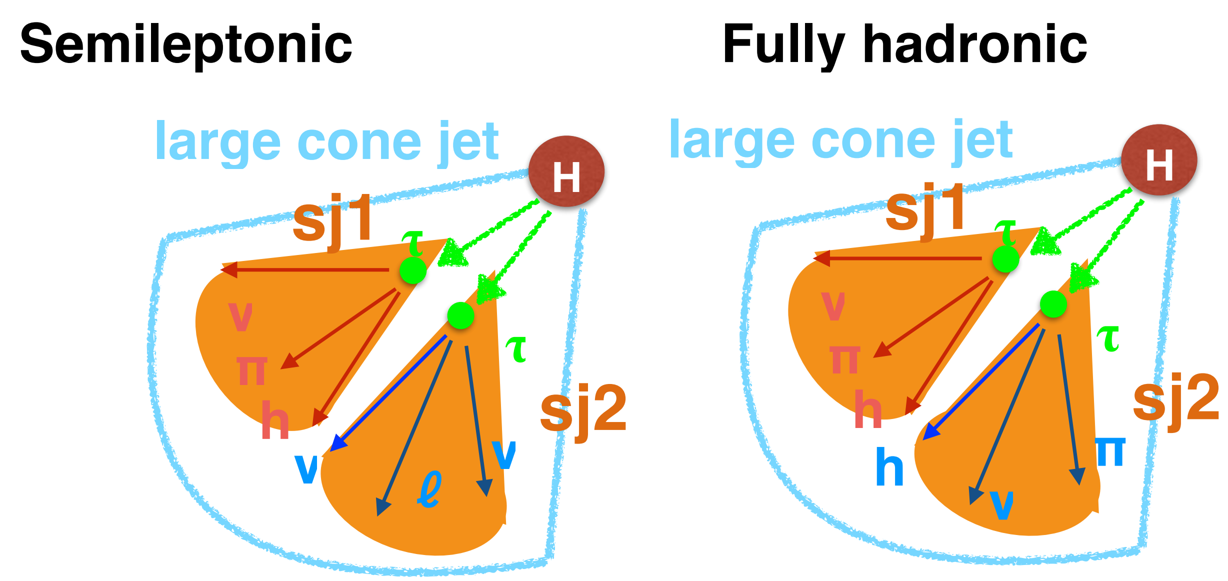


Summary / Outlook

- Hadronic objects are of paramount importance for the CMS physics program
- Wide range of objects identified
  - $\tau \rightarrow \nu$+hadrons
  - Jet flavour (b, c, quark/gluon)
  - Boosted objects (top, W/Z, X → bb)
- Constant struggle to improve current performance and identify more objects
  - Modern machine-learning tools introduced wisely can greatly boost the performance

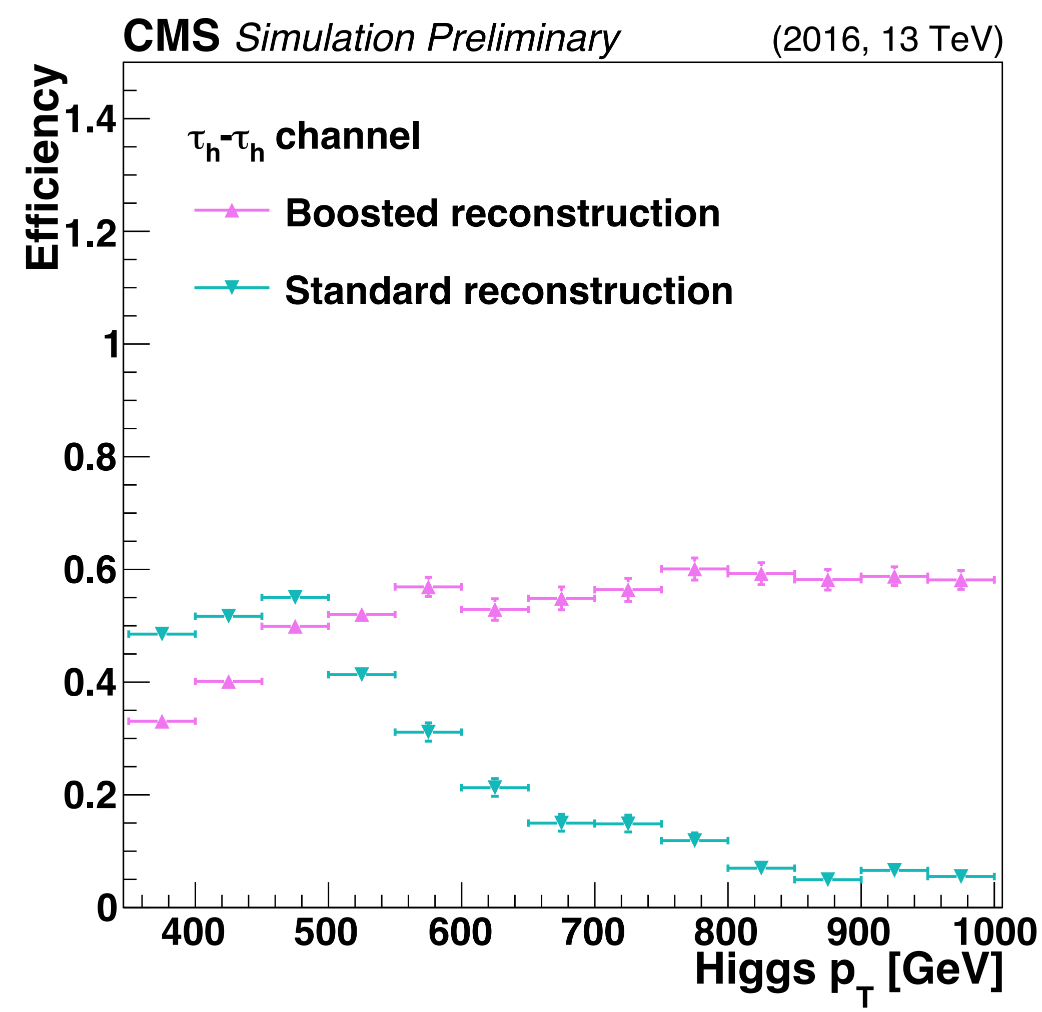
Bonus slides

Boosted tau tagging

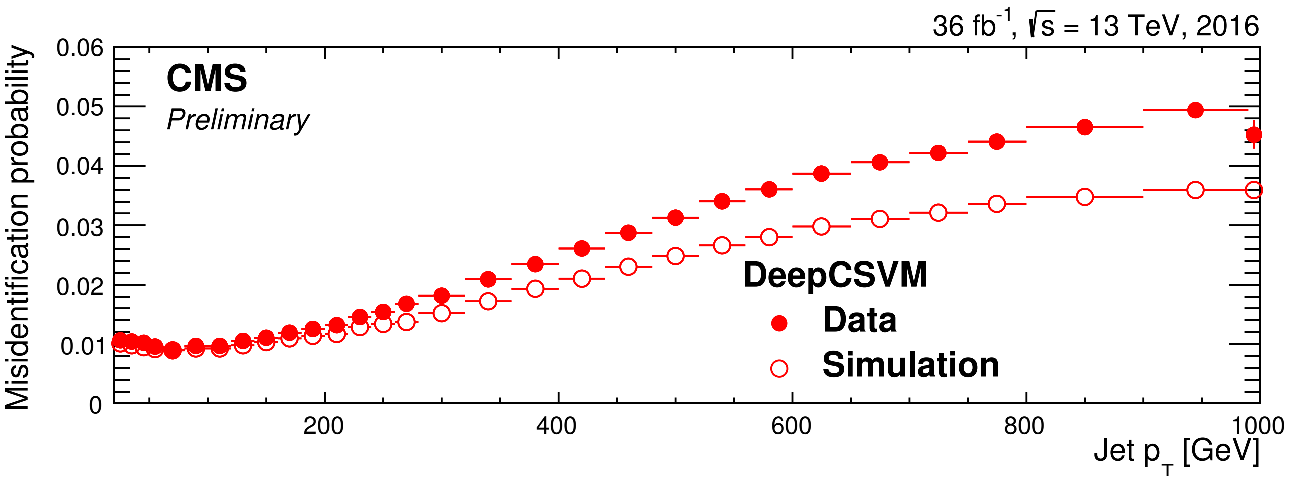


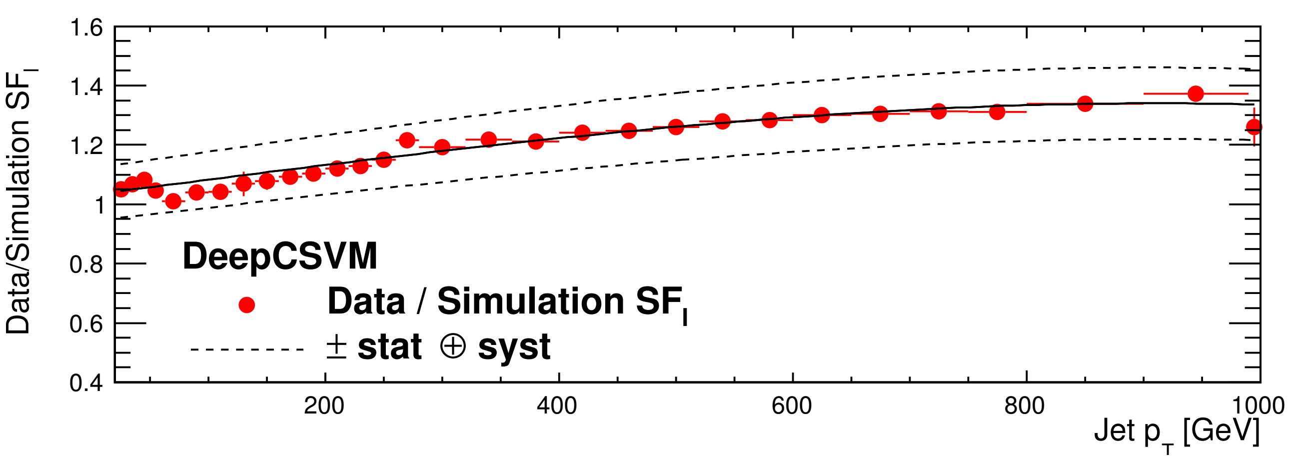
Unified approach for both τ_h and τ_ℓ:

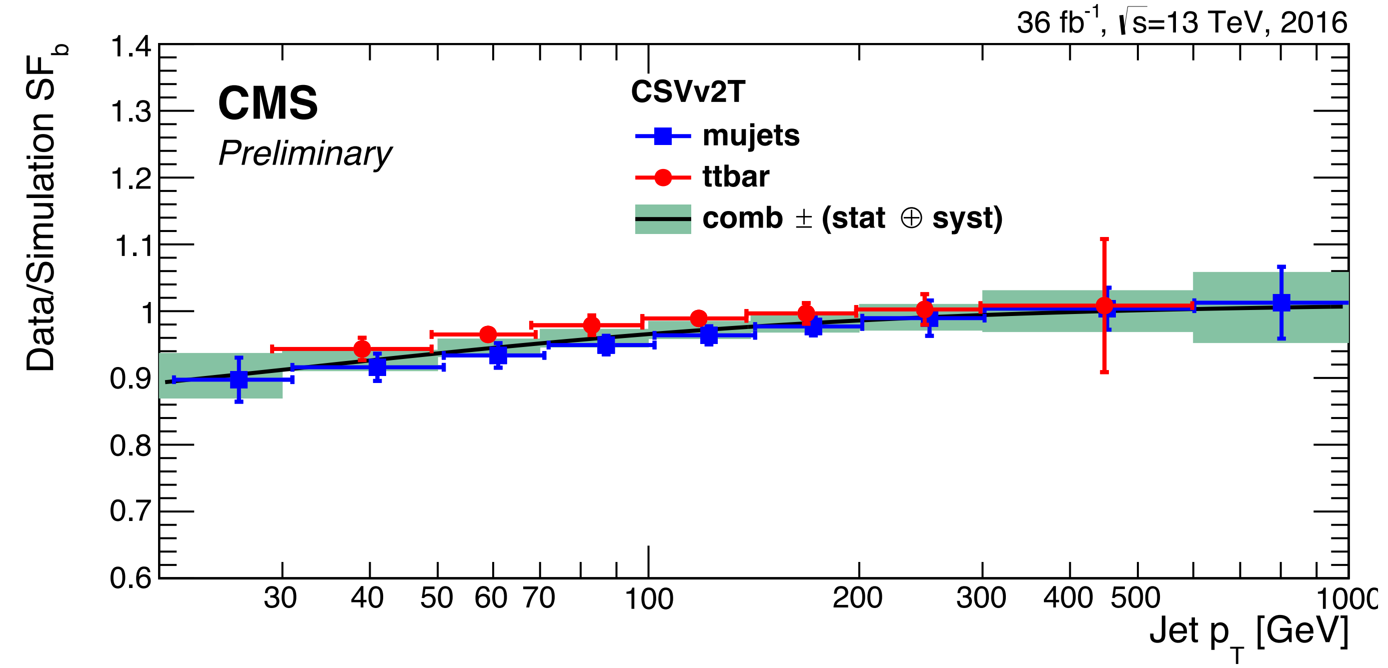
1. Start from AK8 jet
2. Identify sub-jet with p_T > 10 GeV
3. Run standard reconstruction on sub-jet constituents

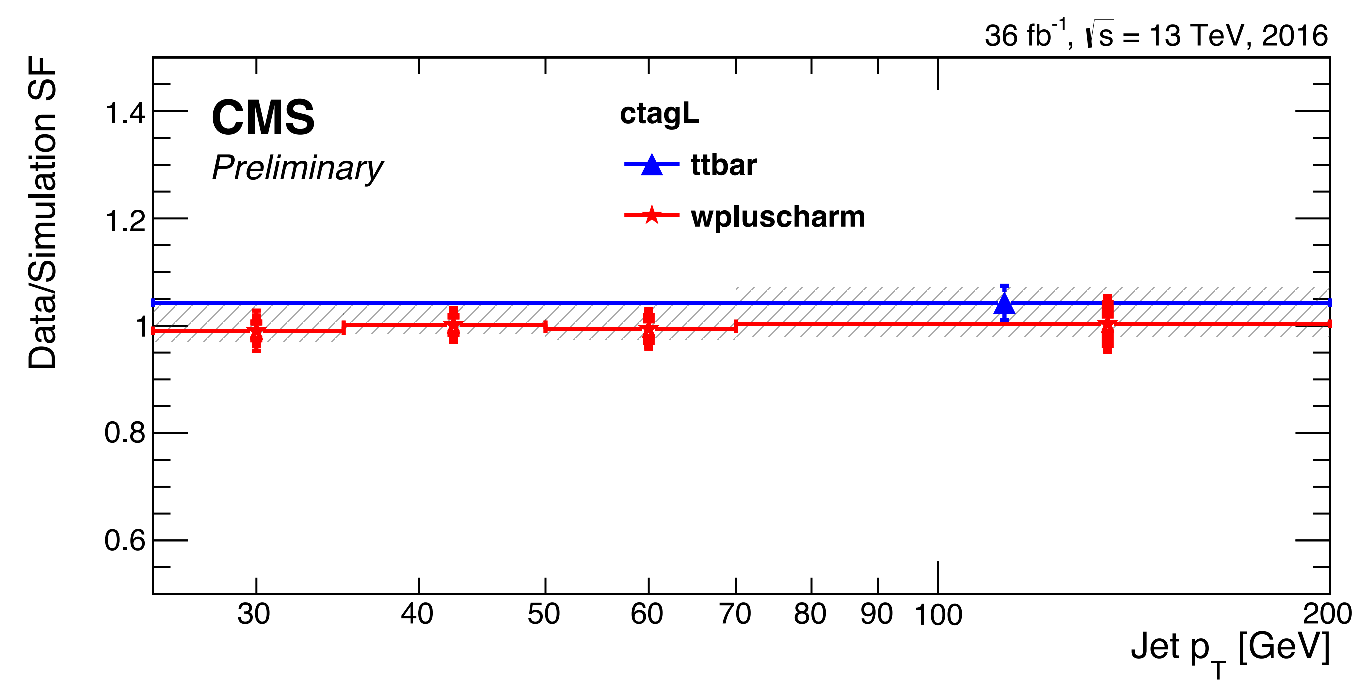


b-tagging performance measurements

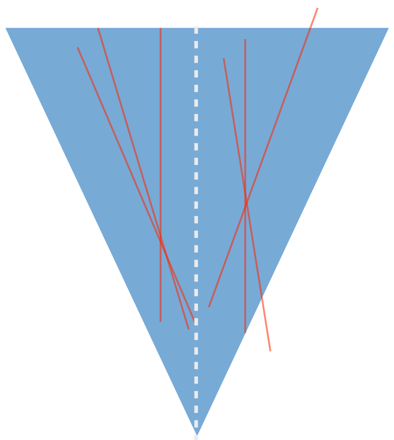




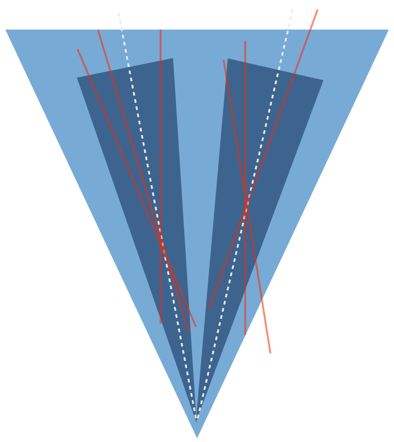




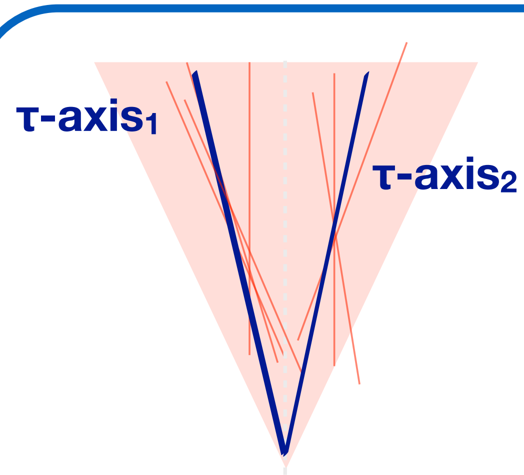
Boosted b-tagging



FatJet: CSVv2 w/o retraining. Custom (relaxed) track and SV association directly on anti-k_T 0.8

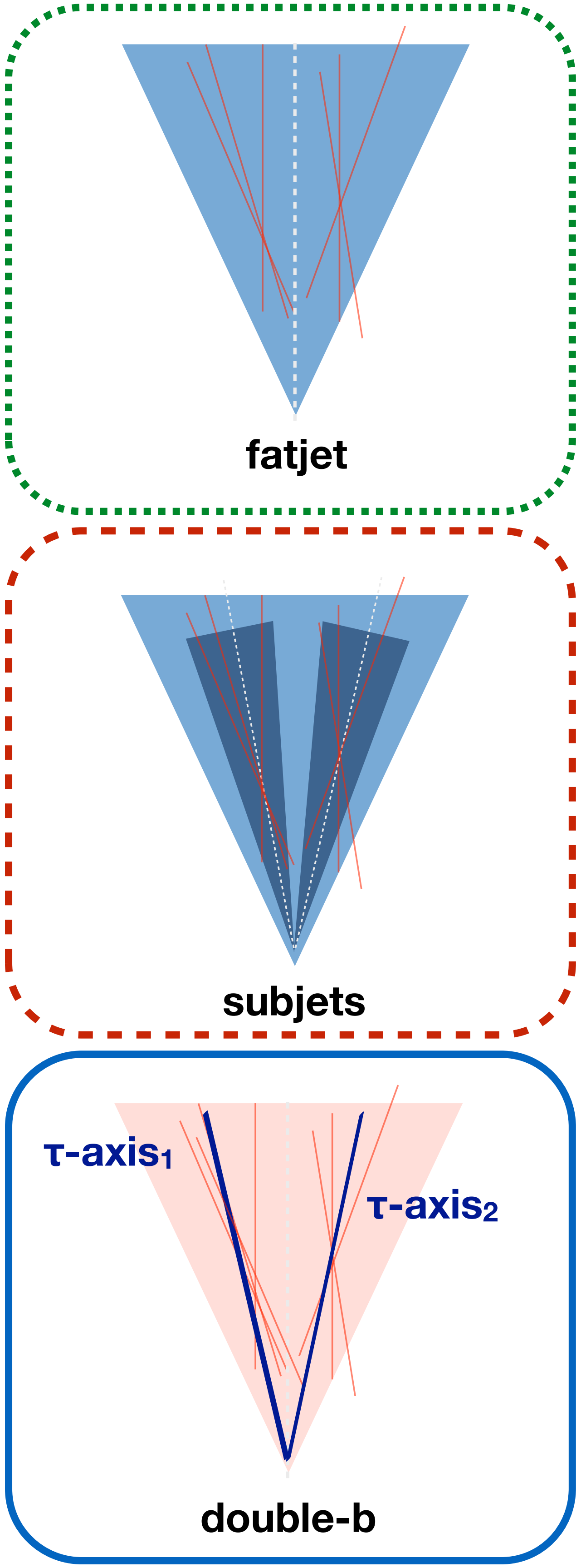


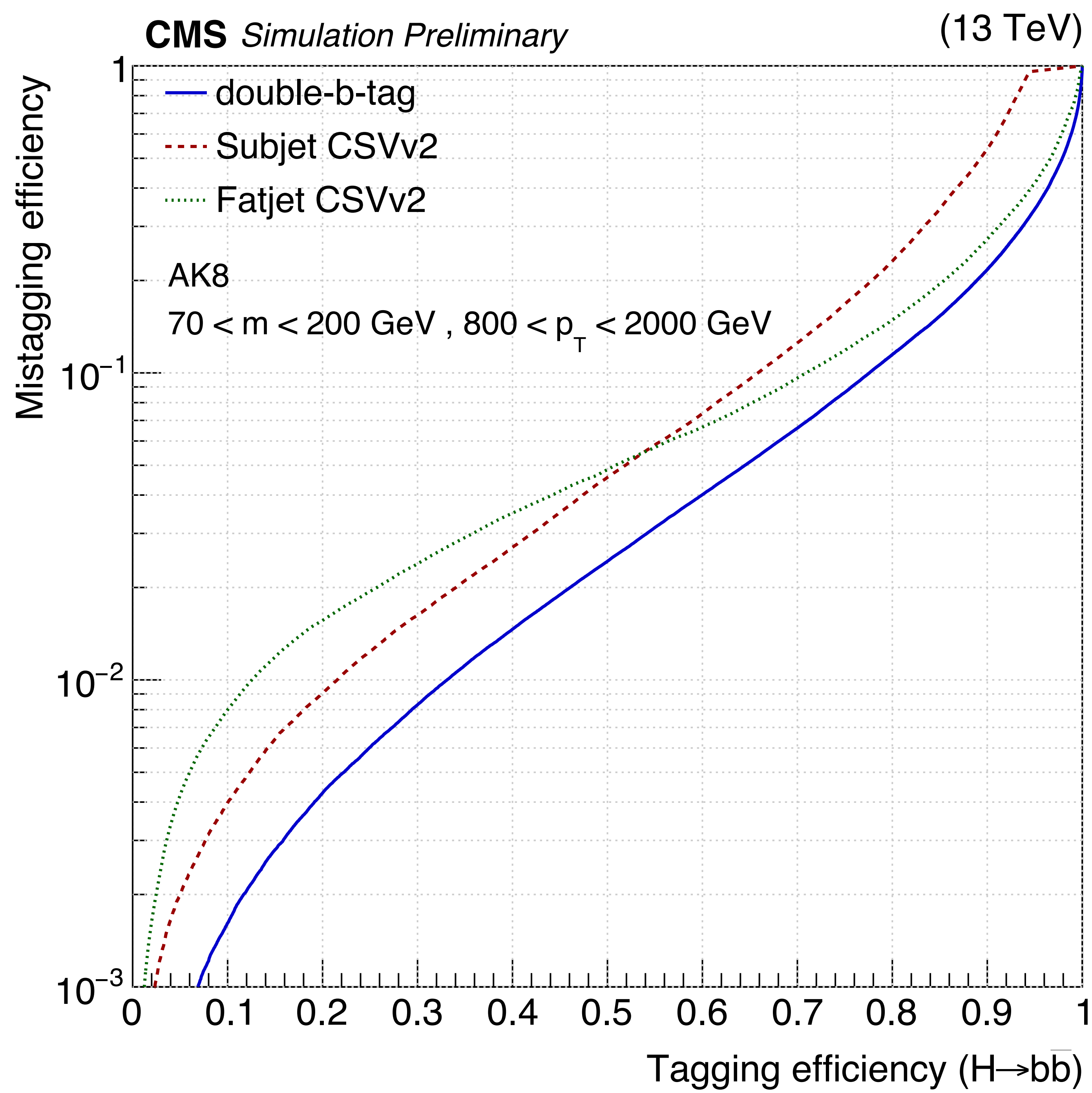
Sub-jet: CSVv2 w/o retraining applied to sub-jets (soft drop, pruned, etc…). Used for boosted top



Double b: dedicated training targeting boosted resonances $X \rightarrow bb$

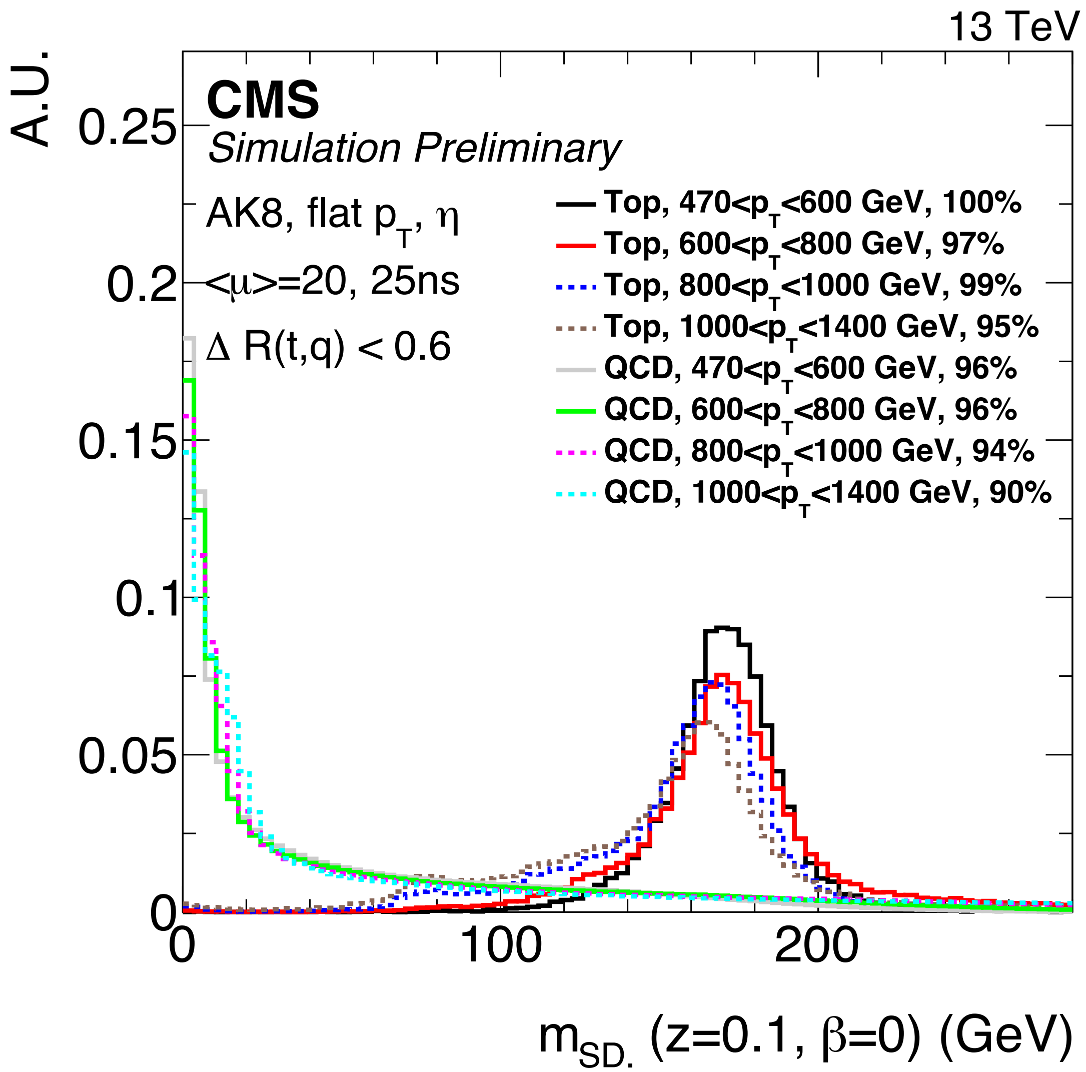
Boosted b-tagging

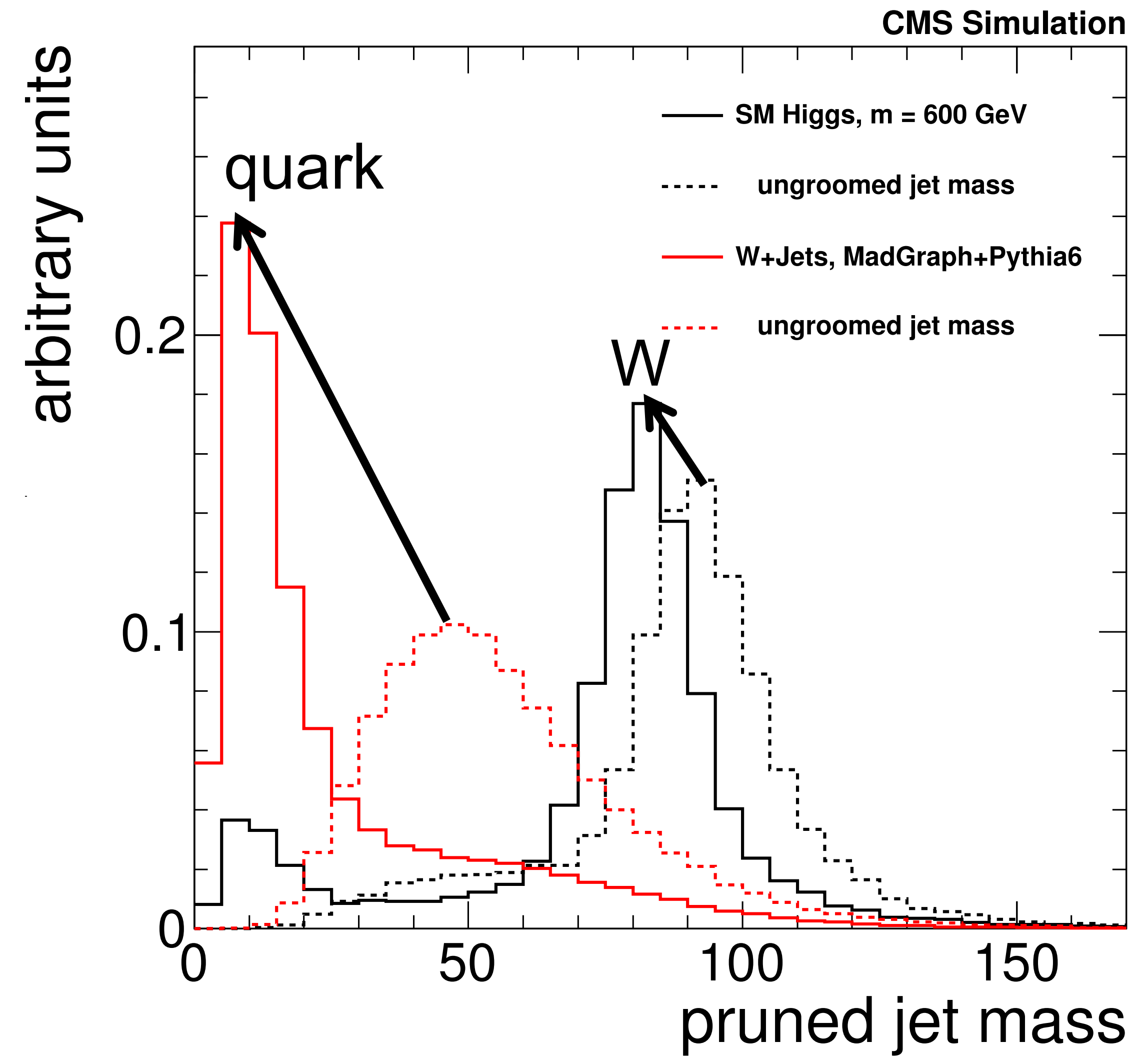




BTV-15-002

SoftDrop (arXiv:1307.0007, arXiv:1402.2657) and pruning (arXiv:0912.0033)





Sub-jettiness

$$\tau_i = \frac{1}{\sum_k p_{T,k} R} \sum_k p_{T,k} \min\left(\Delta R_{1k}, \Delta R_{2k}, \cdots, \Delta R_{ik}\right),$$