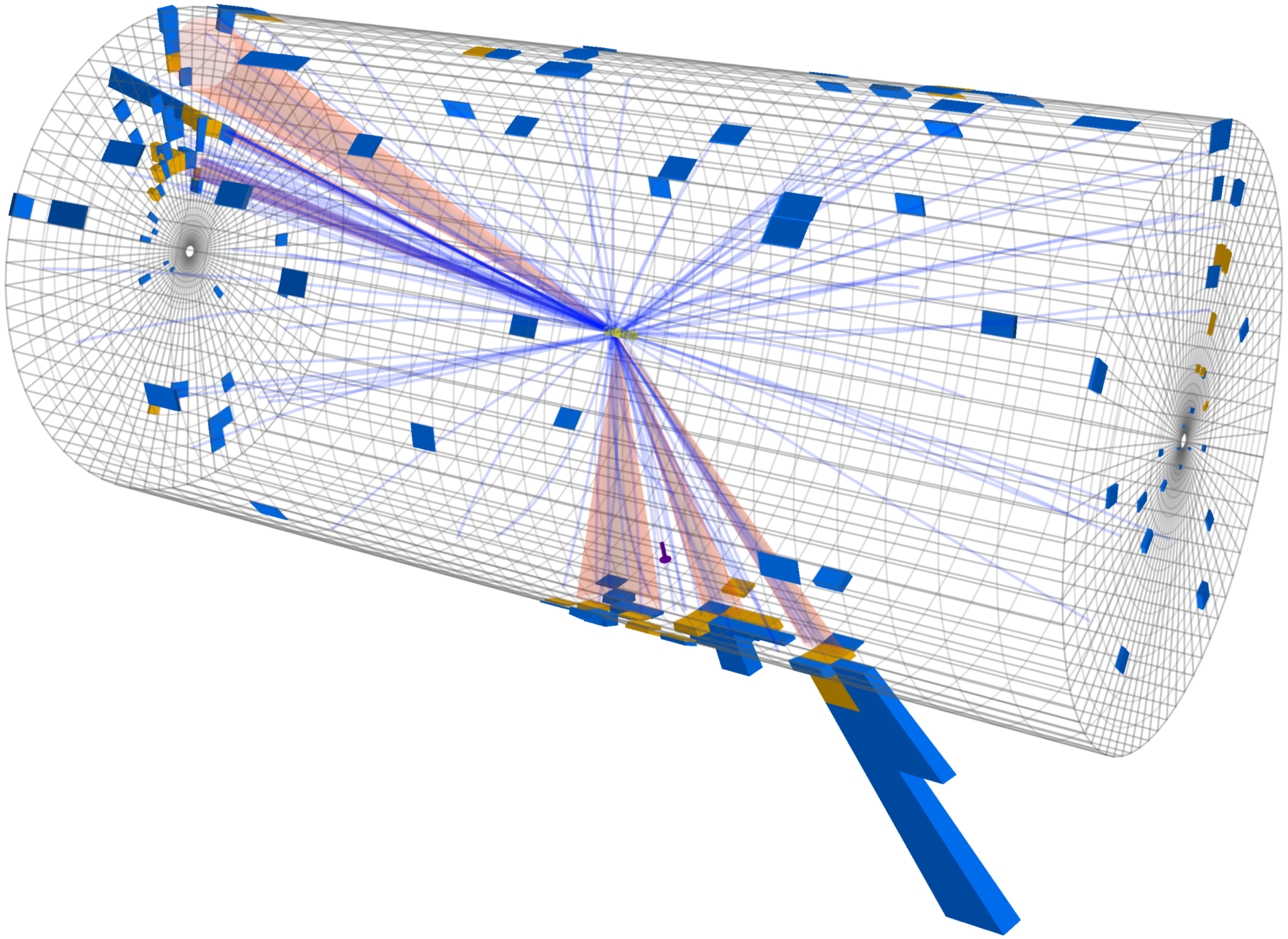
Top Reconstruction & Boosted Top

Experimental Overview



Louise Skinnari (Cornell University)

on behalf of the ATLAS & CMS collaborations

TOP2015, Ischia - September 14-18, 2015

Top & $t\bar{t}$ Reconstruction

- Reconstruct, identify & correctly assign decay products to original top quarks

- **Resolved top decays**
  - Well separated jets
  - Isolated leptons

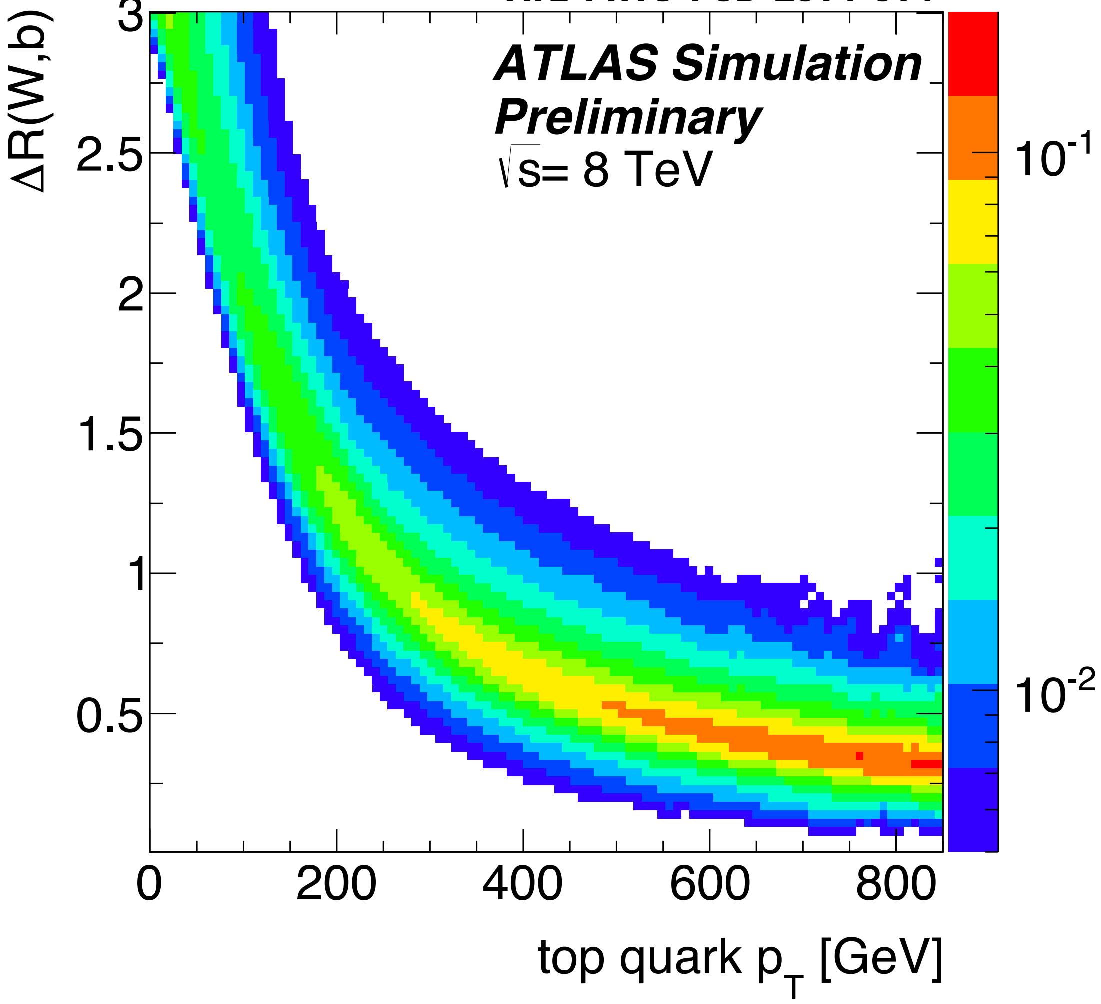
- **Boosted top decays**
  - Overlapping decay products - merged jets
  - Non-isolated leptons

- **Rule of thumb:** $\Delta R \approx 2m/p_{\mathrm{T}}$

  e.g. p_{T}(top) = 350 GeV → ΔR = 1.0

- *MANY different techniques, here showing some examples!*

ATL-PHYS-PUB-2014-014

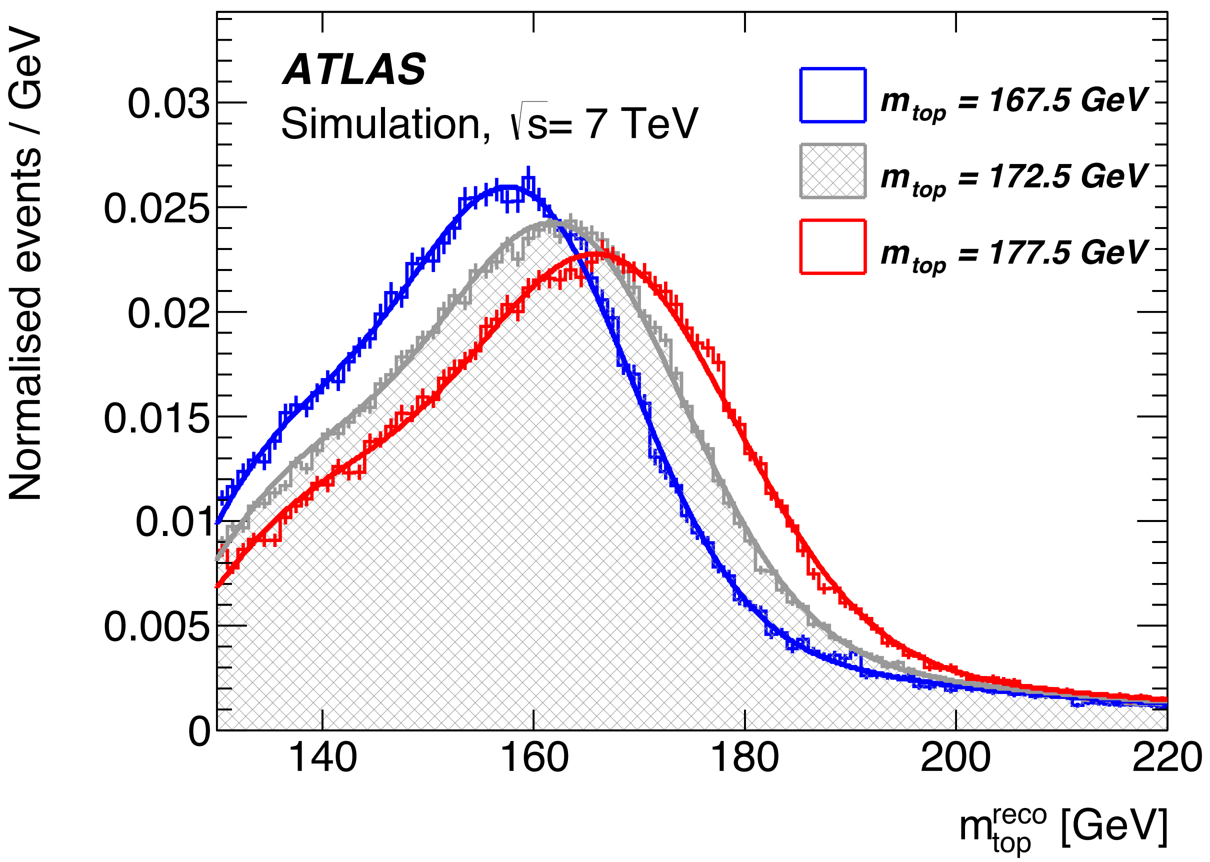


Kinematic Likelihood (ATLAS)

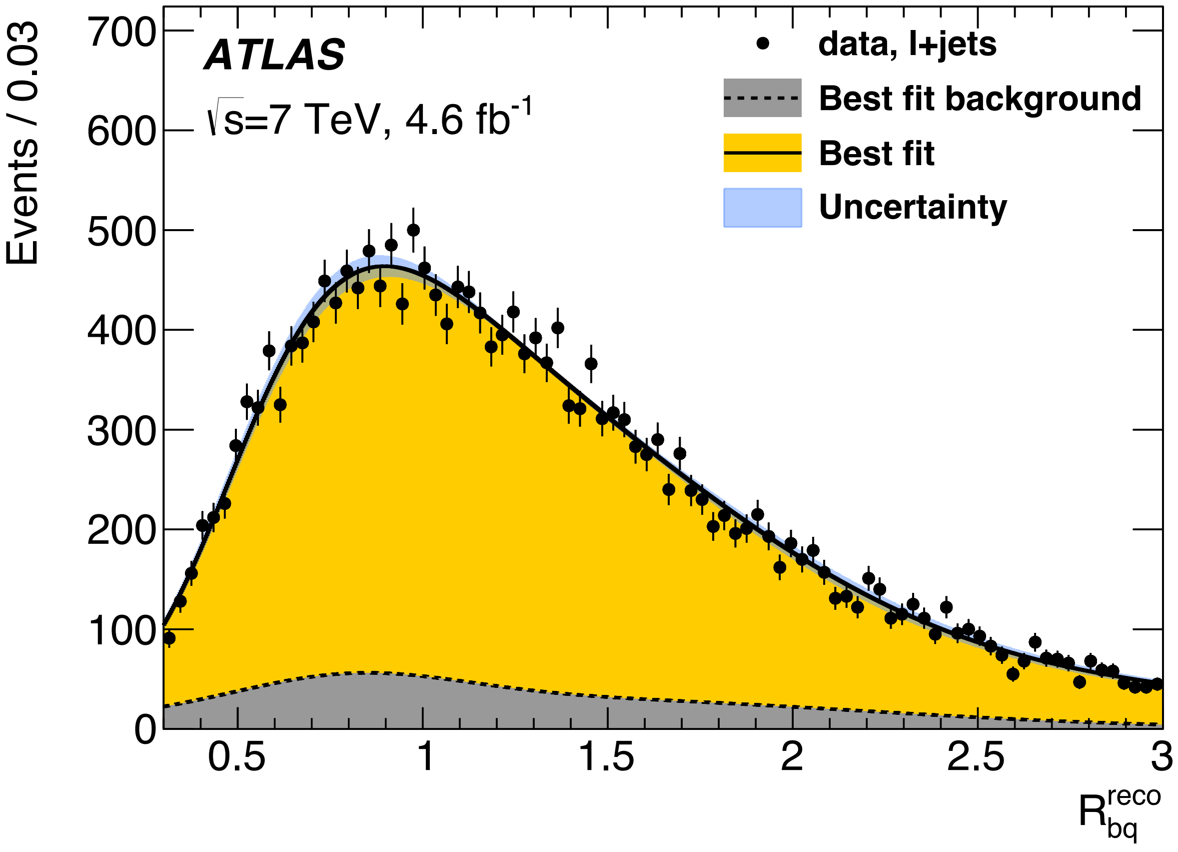
Nucl.Instrum.Meth.
A748 (2014)

- 3D template fit assuming $t\bar{t} \rightarrow$ l+jets topology (KLFitter, standard ATLAS tool)
  - Identify correct jet-parton assignment

- Maximize likelihood, test each permutation of jet-parton association
  - **Breit-Wigner functions:** Constrain dijet/trijet mass to W/t mass
  - **Transfer functions:** Map measured jet energy to energy of final state parton
  - Variable R_{bq} sensitive to relative b/light JES



Eur. Phys. J. C (2015) 75:330



Eur. Phys. J. C (2015) 75:330

Kinematic Likelihood (CMS)

NEW!
see O. Hindrichs' talk

CMS-PAS-TOP-15-005

- $t\bar{t} \rightarrow$ l+jets reconstruction for differential cross section (13 TeV)
  - **p(ν) reconstruction:** $[p(\nu)+p(l)]^2 = m_W^2 \qquad [p(\nu)+p(l)+p(b_l)]^2 = m_t^2$

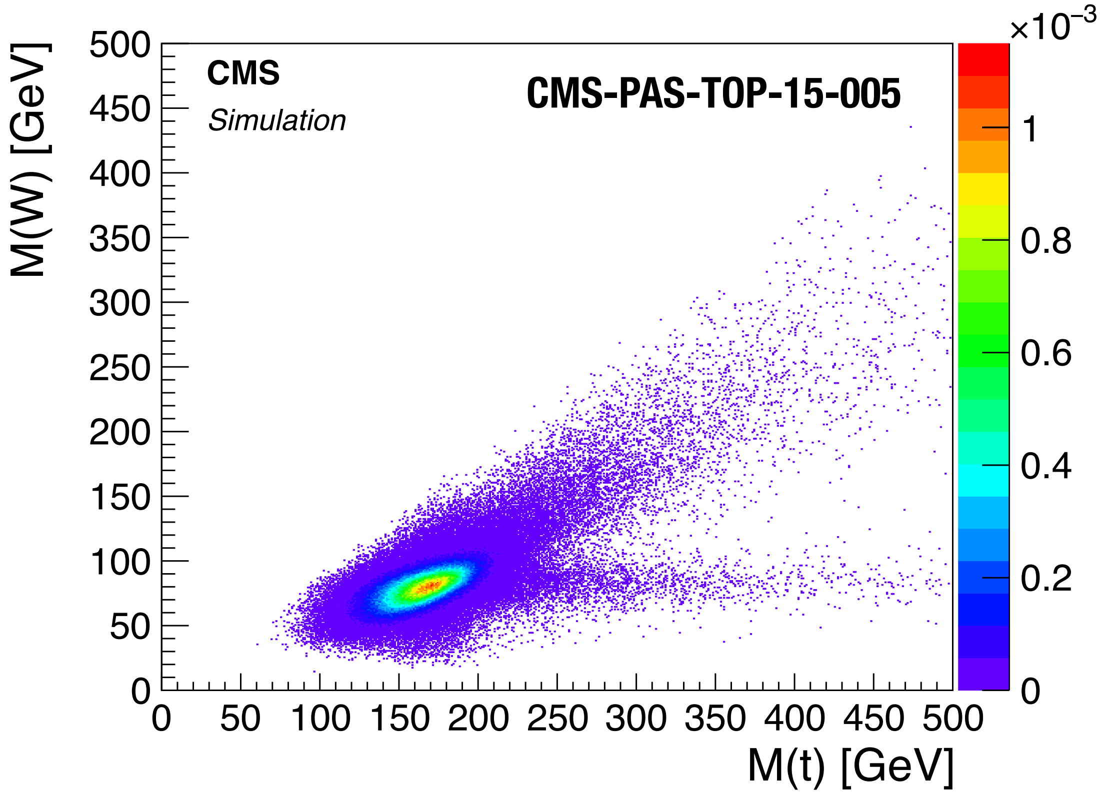
    Solution is ellipse, use point best compatible with $\not{E}_T \rightarrow$ improved p_T (ν)
  - **$t\bar{t}$ reconstruction:** Likelihood for most probable quark-to-jet assignment

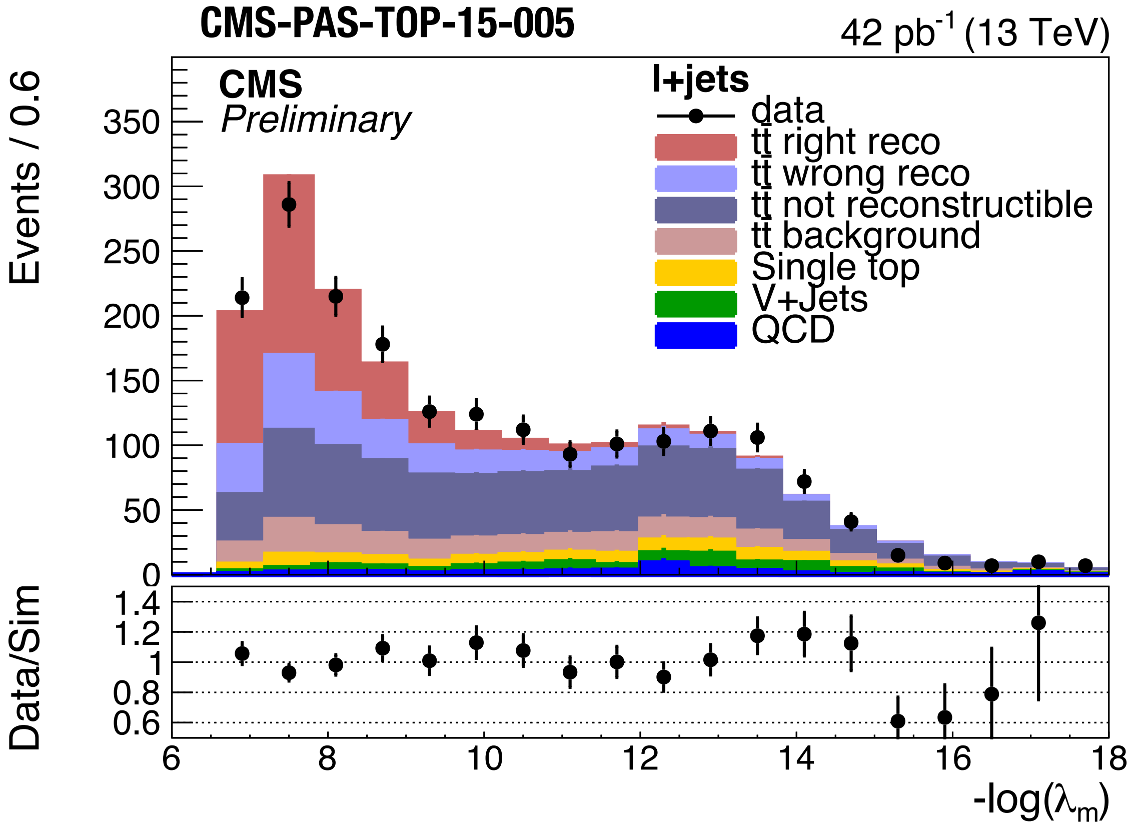
$$-log(\lambda) = -log(\lambda_m) - log(\lambda_\nu)$$
$$\lambda_m = P_m(m_2, m_3)$$
$$\lambda_\nu = P_\nu(D_{\nu,min})$$

2D probability for m_{inv} of jets tested as W (m_2) & hadronic top (m_3)

Probability of ν reconstruction for correctly selected b-jet

- ~60% reconstruction efficiency





Pseudo-Top Reconstruction

JHEP 06 (2015) 100
arXiv:1505.04480

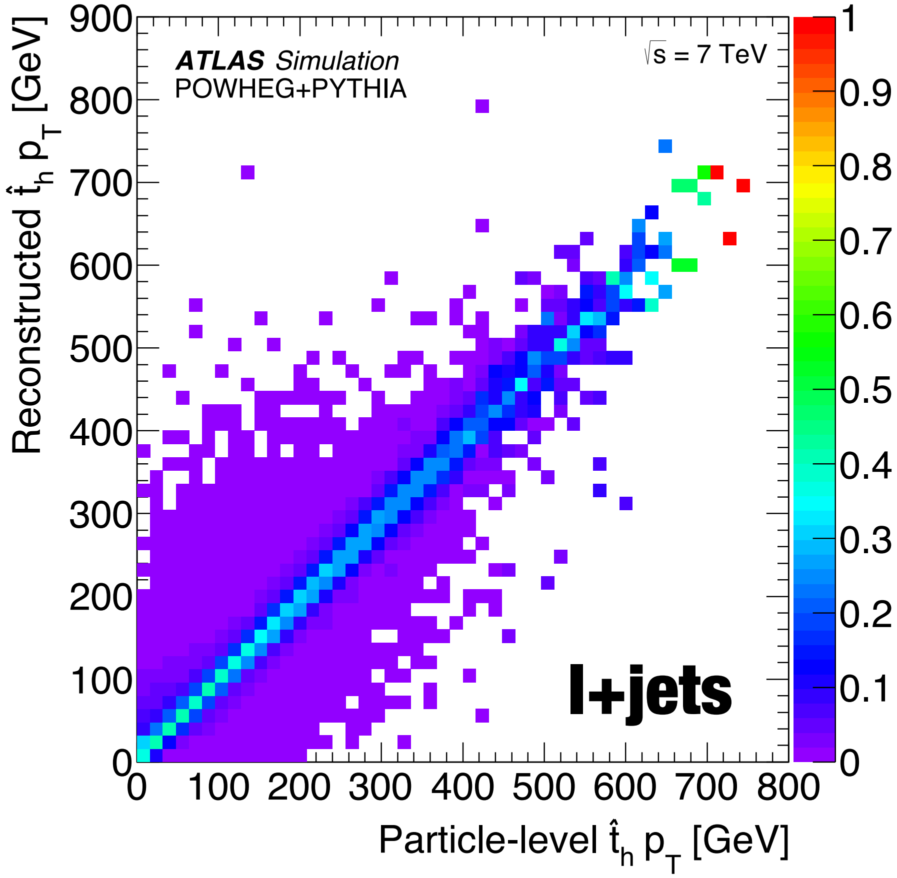
- Minimize model-dependence in differential cross section results
  - Allow QCD precision tests in top quark sector
- Top-quark proxy at stable-particle level

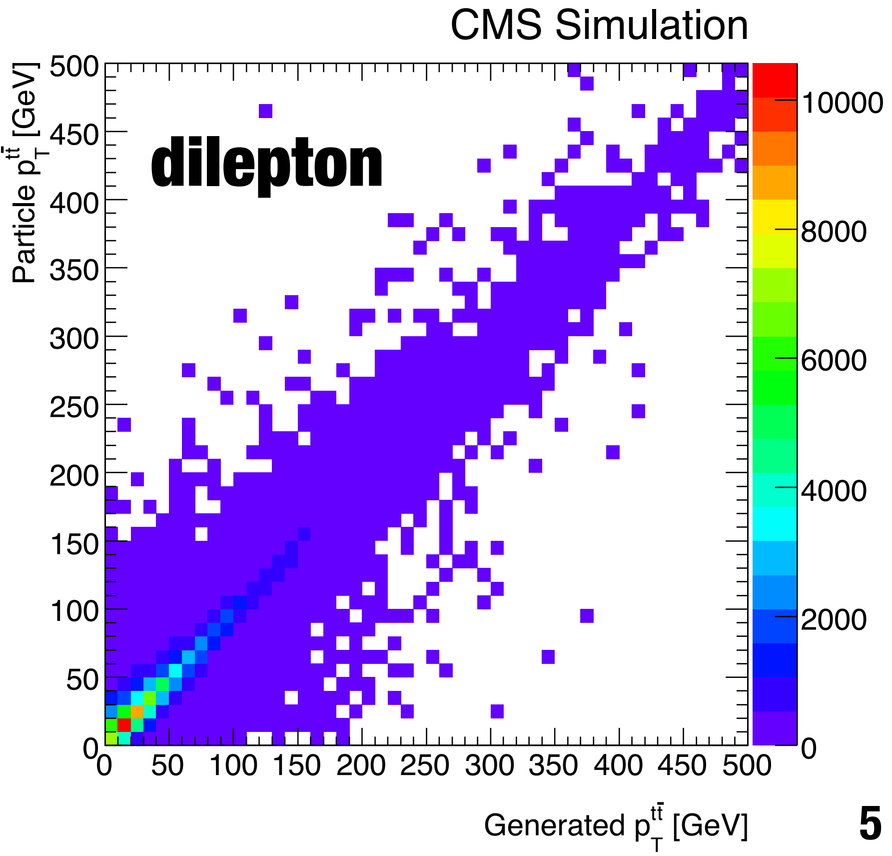
- **W_{lep}:** e/μ + $\not{E}_T$, solve for $p_{z,\nu}$ assuming m_W
- **$\hat{t}_{lep}$:** W_{lep} + closest b-jet
- **W_{had}:** two other highest-p_T jets
- **$\hat{t}_{had}$:** W_{had} + remaining b-jet

l+jets

- Consider two leading-p_T neutrinos
- **Pseudo $W^\pm$** from ($\nu_{1/2}$, $l^\pm$) pair that minimize: $|m_{W,1} - m_{W,PDG}|+|m_{W,2} - m_{W,PDG}|$
- **Pseudo tops** from ($b_{1/2}$, $W^\pm$) pair that minimize: $|m_{t,1} - m_{t,PDG}|+|m_{t,2} - m_{t,PDG}|$

dilepton



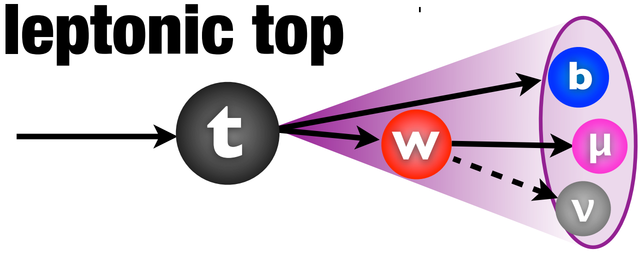


Boosted Top Quarks

Many ATLAS/CMS searches with boosted tops:
$t\bar{t}$ / tb / tH resonances, stop quarks, vector-like quarks, ...

- Why / how?
  - Test predictions of high-p_T top production
  - Probe new physics -- many models predict new particles at the TeV scale
  - Collimated decay products → special reconstruction techniques

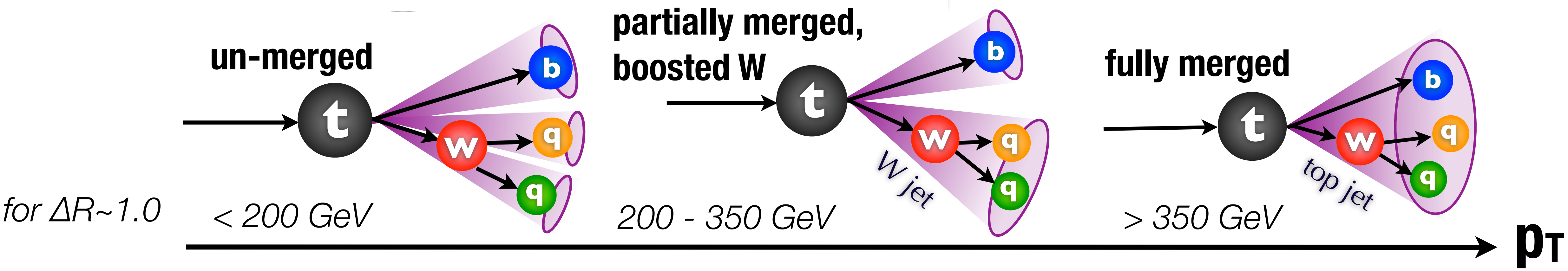
boosted leptonic top



- Lepton close to b-jet
- Standard lepton isolation suboptimal → alternative definitions

boosted hadronic top

- Cluster decay products in single large-R jet
- Jet substructure to distinguish signal from QCD background
- Jet grooming to remove soft radiation


un-merged

partially merged, boosted W

fully merged

for ΔR~1.0

< 200 GeV

200 - 350 GeV

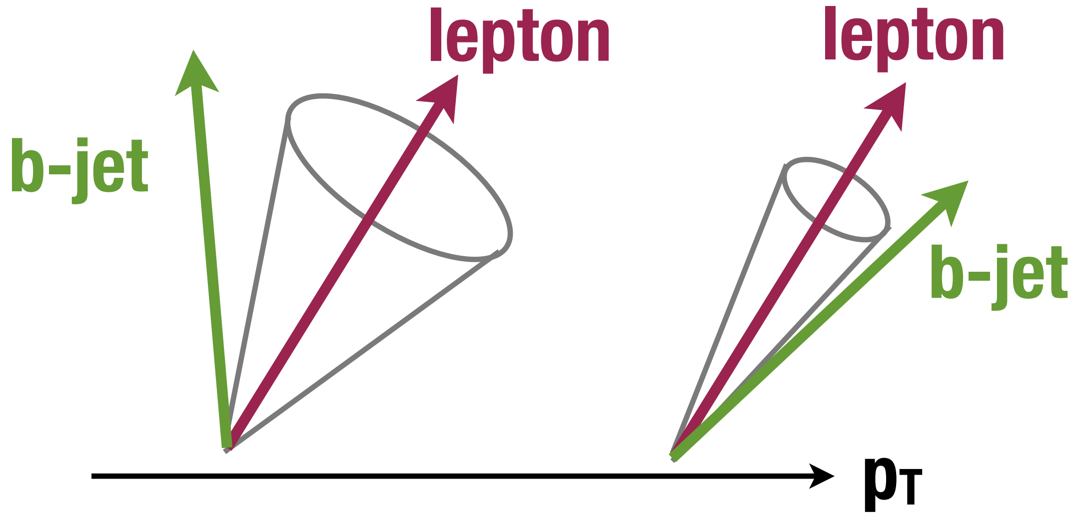
> 350 GeV

Boosted Tops -- Leptonic Decays

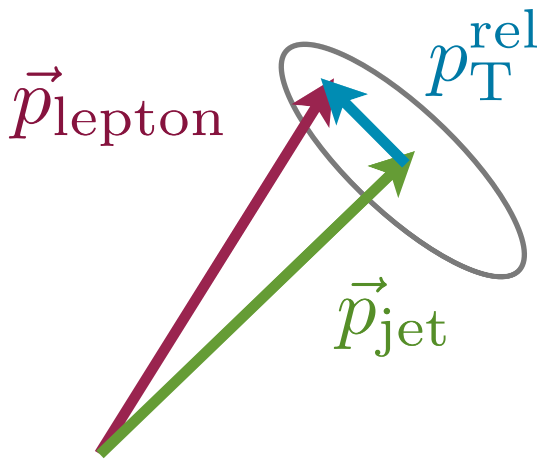
- Shrinking isolation cone (ATLAS)

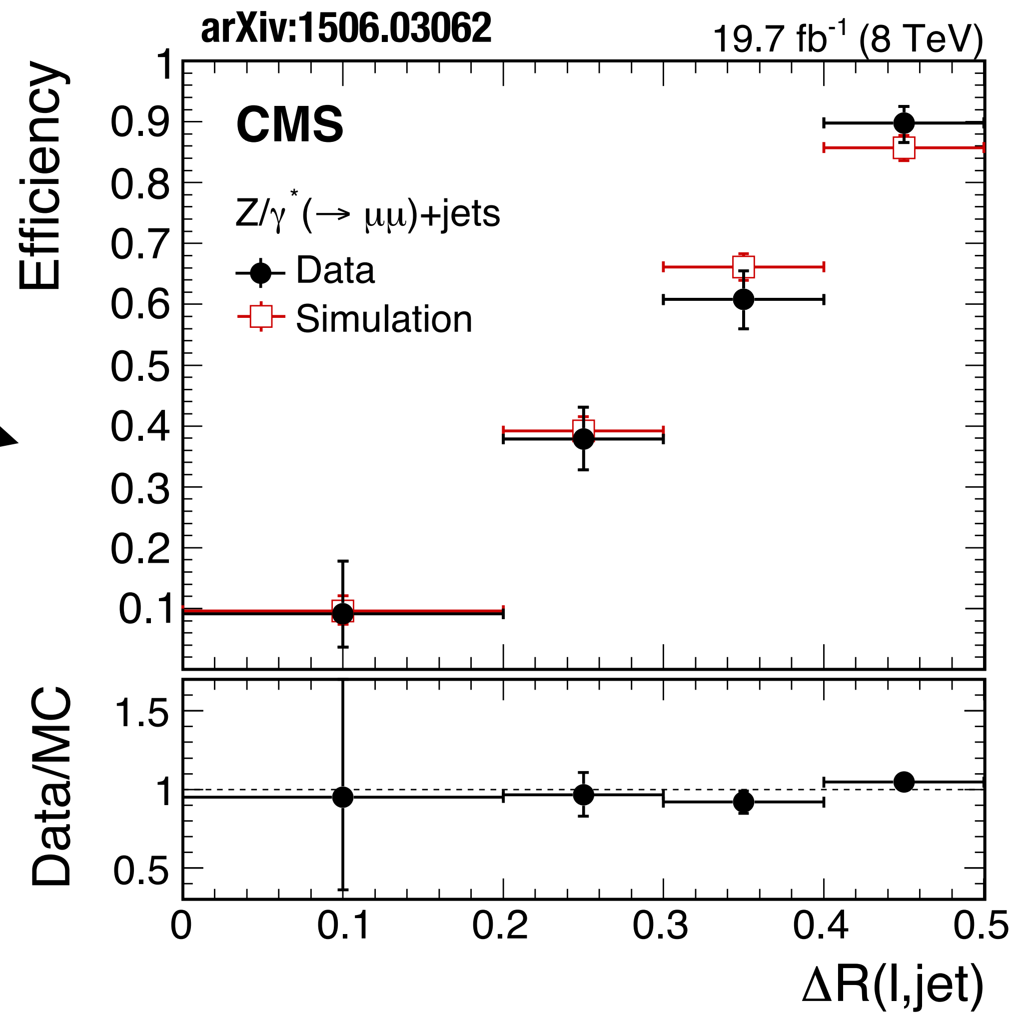
$$I_{mini} = \sum_{\mathrm{tracks}} p_{\mathrm{T}}^{\mathrm{track}} / p_{\mathrm{T}}^{l}, \;\; \Delta R(l, \mathrm{track}) < K_T / p_{\mathrm{T}}^{l}$$

K_T = 10 GeV →
ΔR = 0.4 (0.1) at 25 (100) GeV
Require $I_{mini} < 0.05$



- Kinematic cuts (CMS)
  - ΔR(lepton, jet) separation
  - Component of lepton p_{T} transverse to jet axis





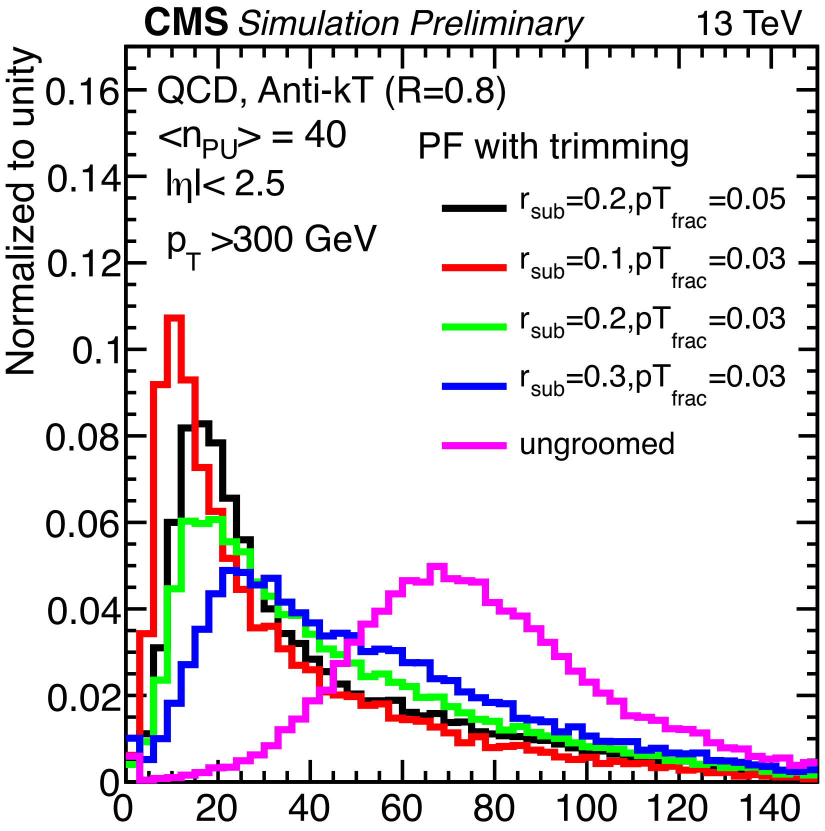
Jet Grooming

PRL 100 (2008) 242001
JHEP 1002 (2010) 084
PRD 81 (2010) 094023
JHEP 1405 (2014) 146

- Techniques to reduce contamination from ISR, underlying event & pileup
    - **Filtering:** *Recluster, keep N hardest subjets*
    - **Trimming:** *Recluster, keep subjets with* $p_{T,i} / p_{T,jet} > x$
    - **Pruning:** *Reject soft & large angle constituents as part of jet algorithm*
    - **Soft drop:** *Remove wide-angle soft radiation*

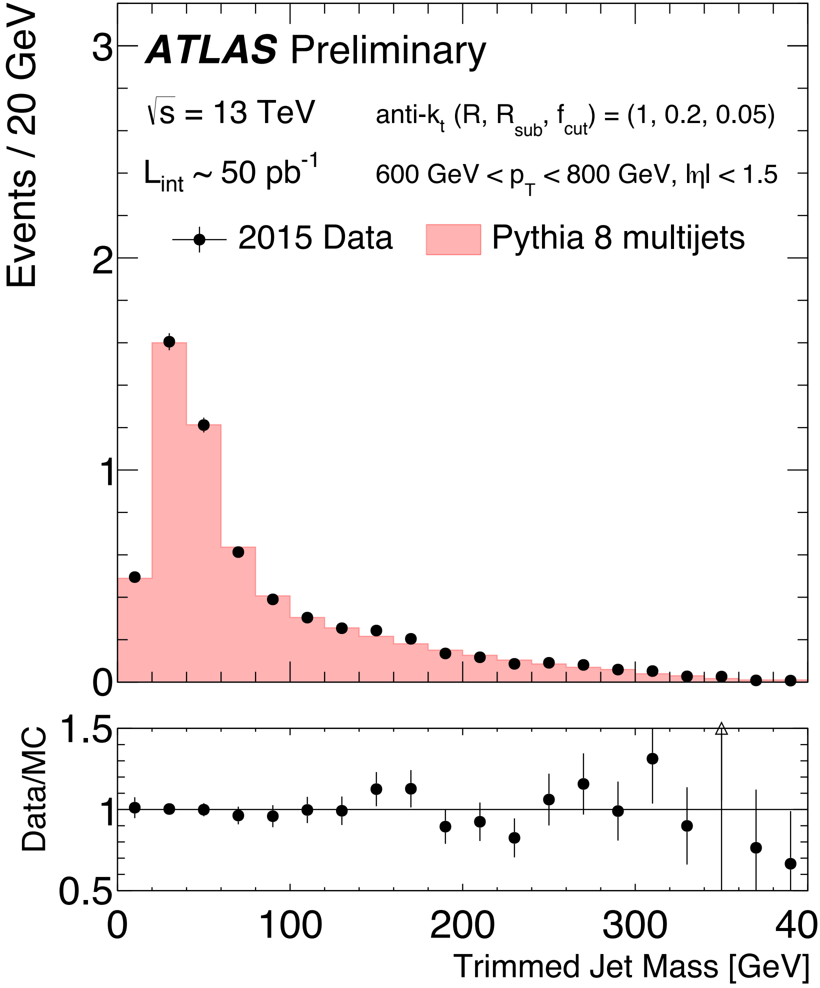
$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\rm cut} \left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$$

tunable variables

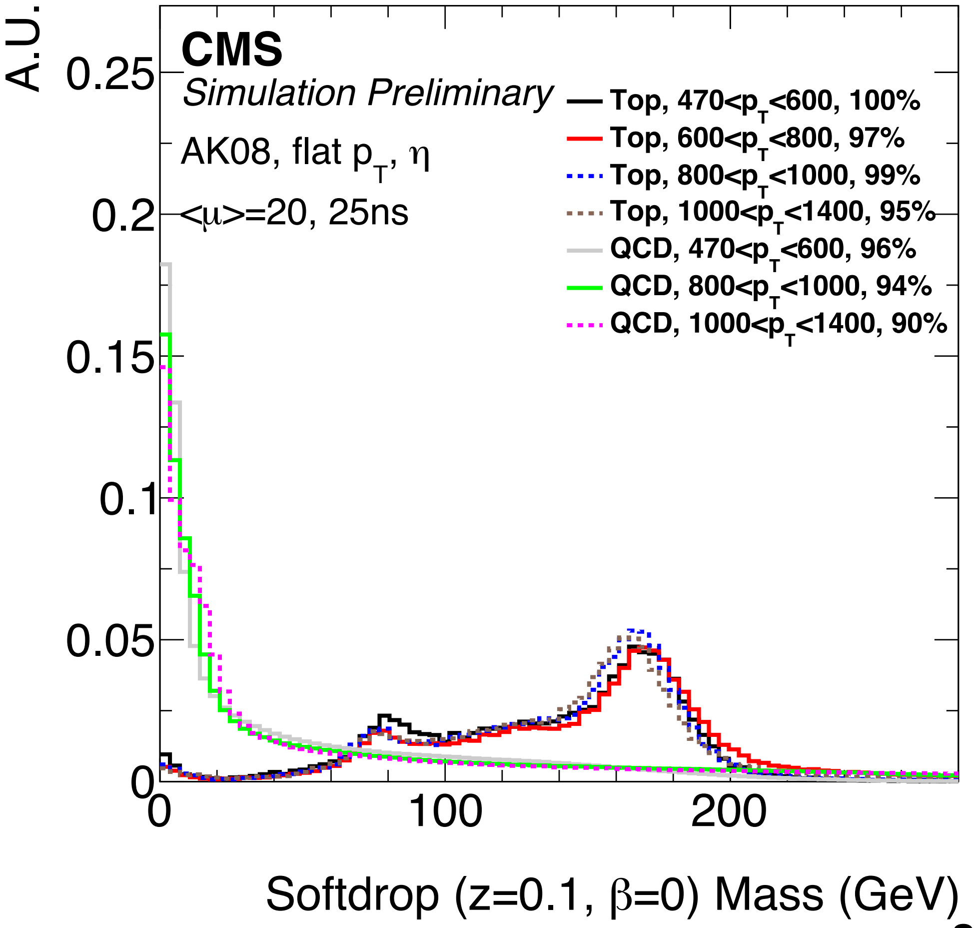


CMS-JME-14-001

ATLAS-CONF-2015-035



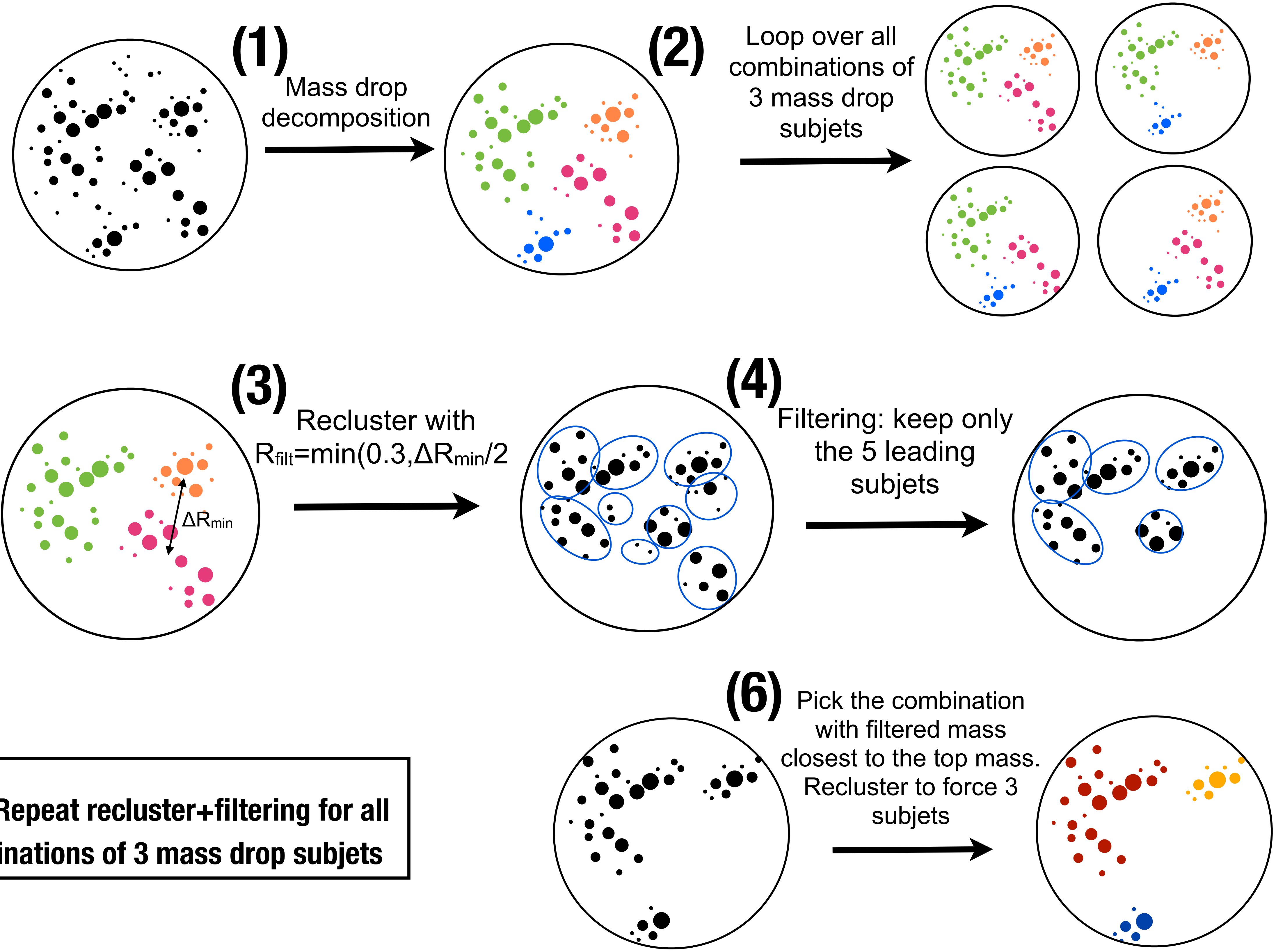
CMS-DP-2015-043



HEP Top Tagger

JHEP 1010 (2010) 078

Start from R=1.5 Cambridge-Aachen jets (used for $p_T > 200$ GeV)



(1) Mass drop decomposition

(2) Loop over all combinations of 3 mass drop subjets

(3) Recluster with $R_{filt}=\min(0.3,\Delta R_{min}/2)$

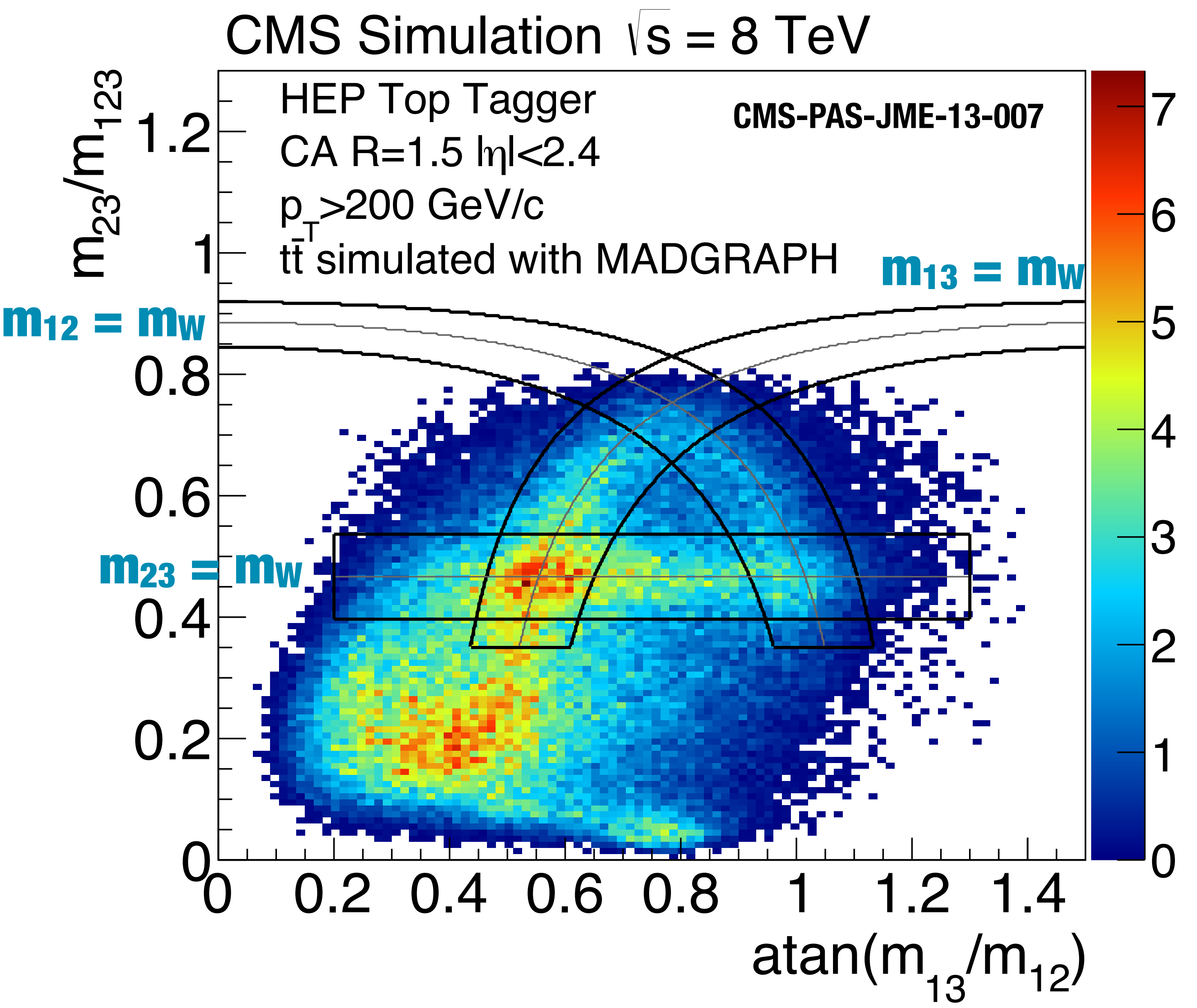
(4) Filtering: keep only the 5 leading subjets

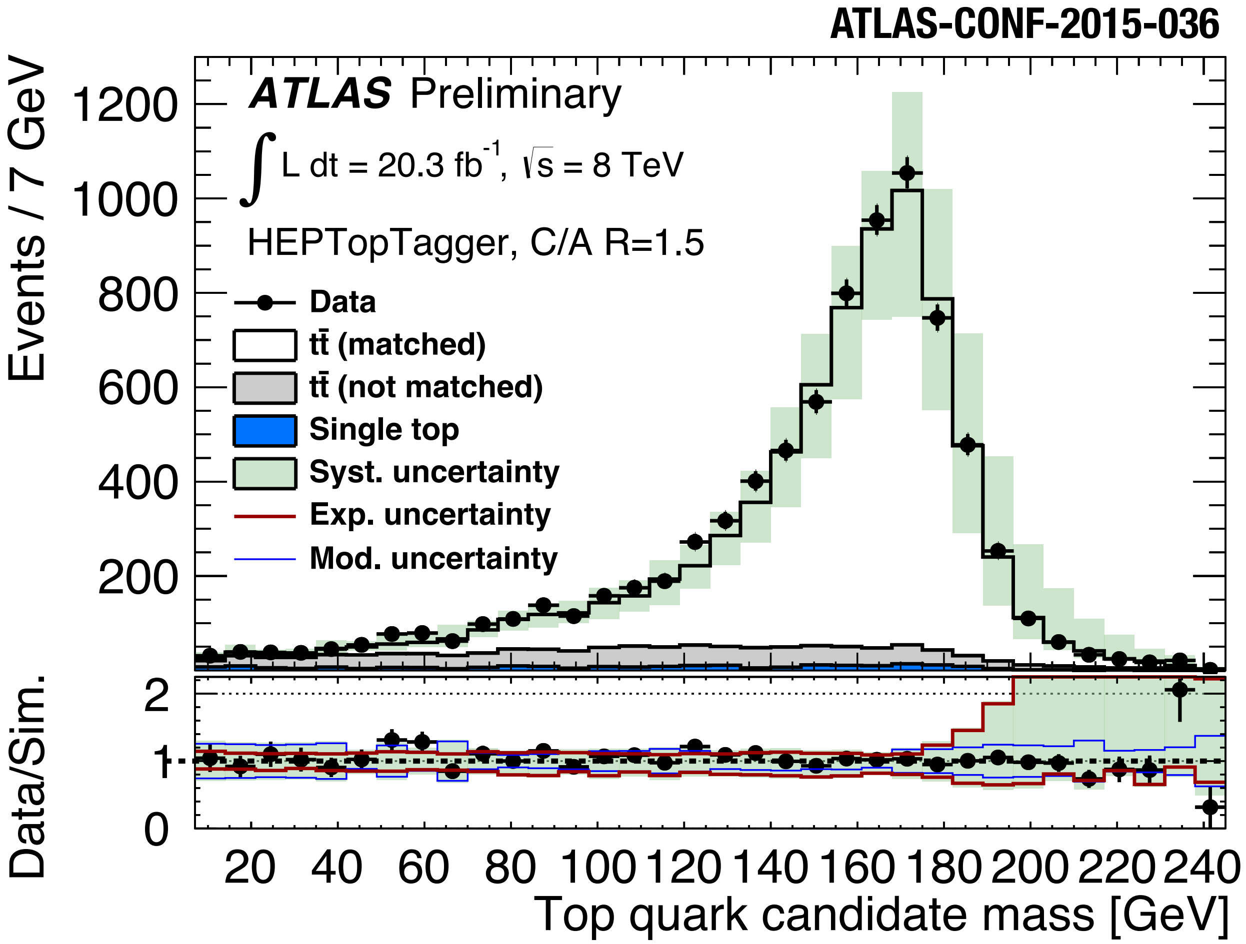
(5) Repeat recluster+filtering for all combinations of 3 mass drop subjets

(6) Pick the combination with filtered mass closest to the top mass. Recluster to force 3 subjets

HEP Top Tagger

- 2D distribution based on ratio of subjet pairwise masses
  - m_{123} = "top jet mass"
  - 1-2-3: subjets p_T ordered
- **W mass selection**





- **Top mass selection**
  140 < m_{123} < 210-250 GeV

Used by e.g. all-hadronic $t\bar{t}$ resonance searches at low boost

JHEP 01 (2013) 116 (ATLAS)
arXiv:1506.03062 (CMS)

CMS Top Tagger

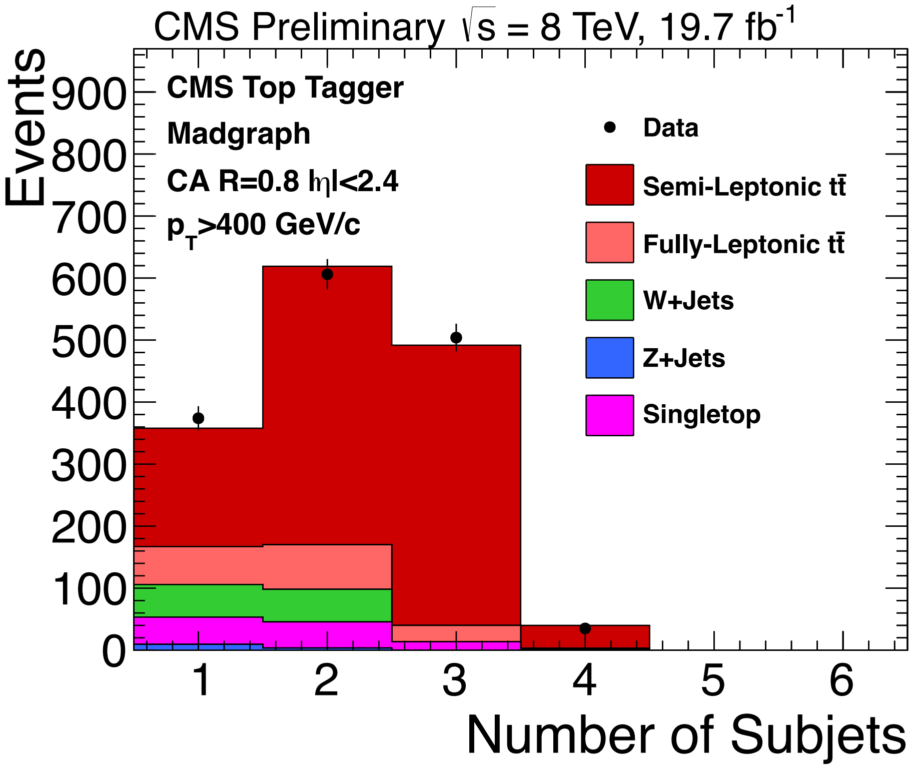
CMS-PAS-JME-13-007

- Decluster jets into subjets, removing soft & wide-angle clusters
  - **Adjacency:** $\Delta R(A,B) > 0.4 - 0.004 \times p_T^{input}$
  - **Softness:** $p_T^{subjet} > 0.05 \times p_T^{hard\ jet}$

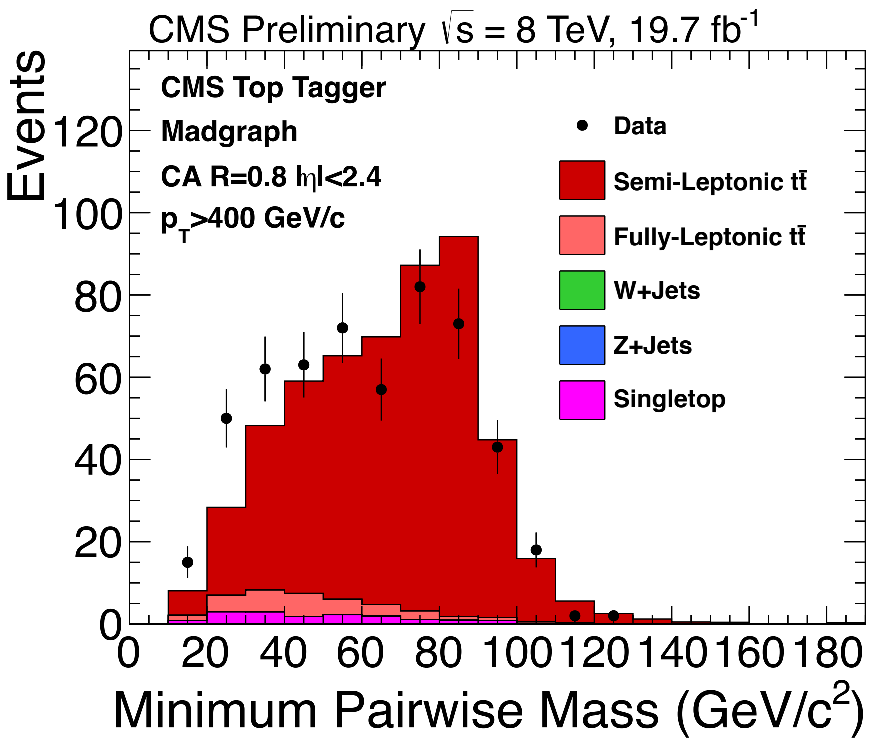
R=0.8 jets → used for p_T > 400 GeV

Require:

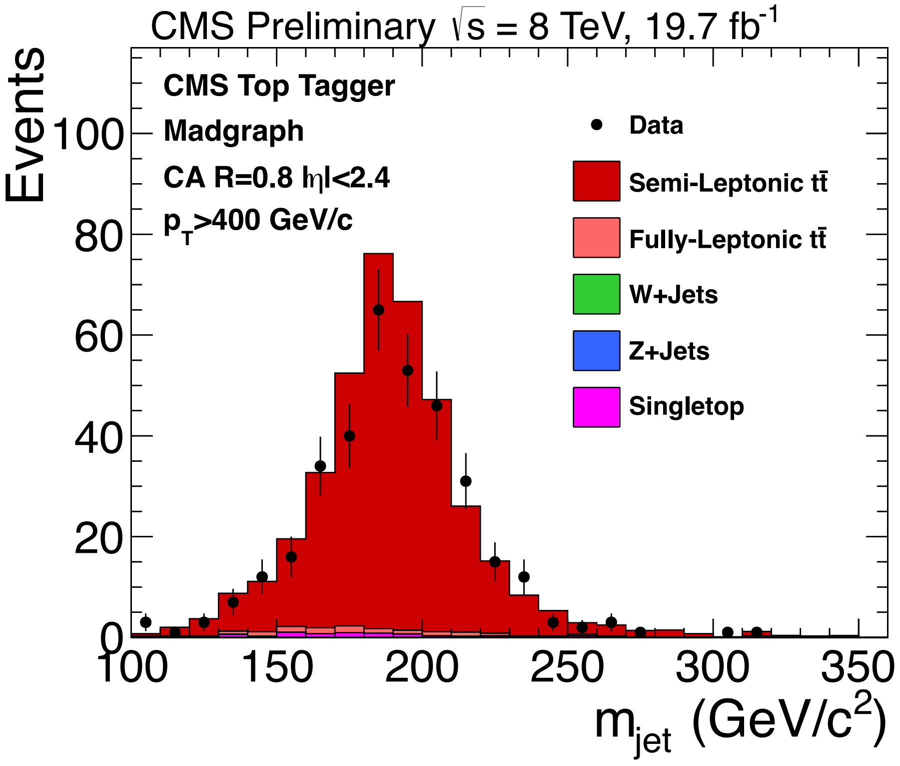
≥ 3 subjets



Min. pairwise subjet mass > 50 GeV



140 < m_{jet} < 250 GeV

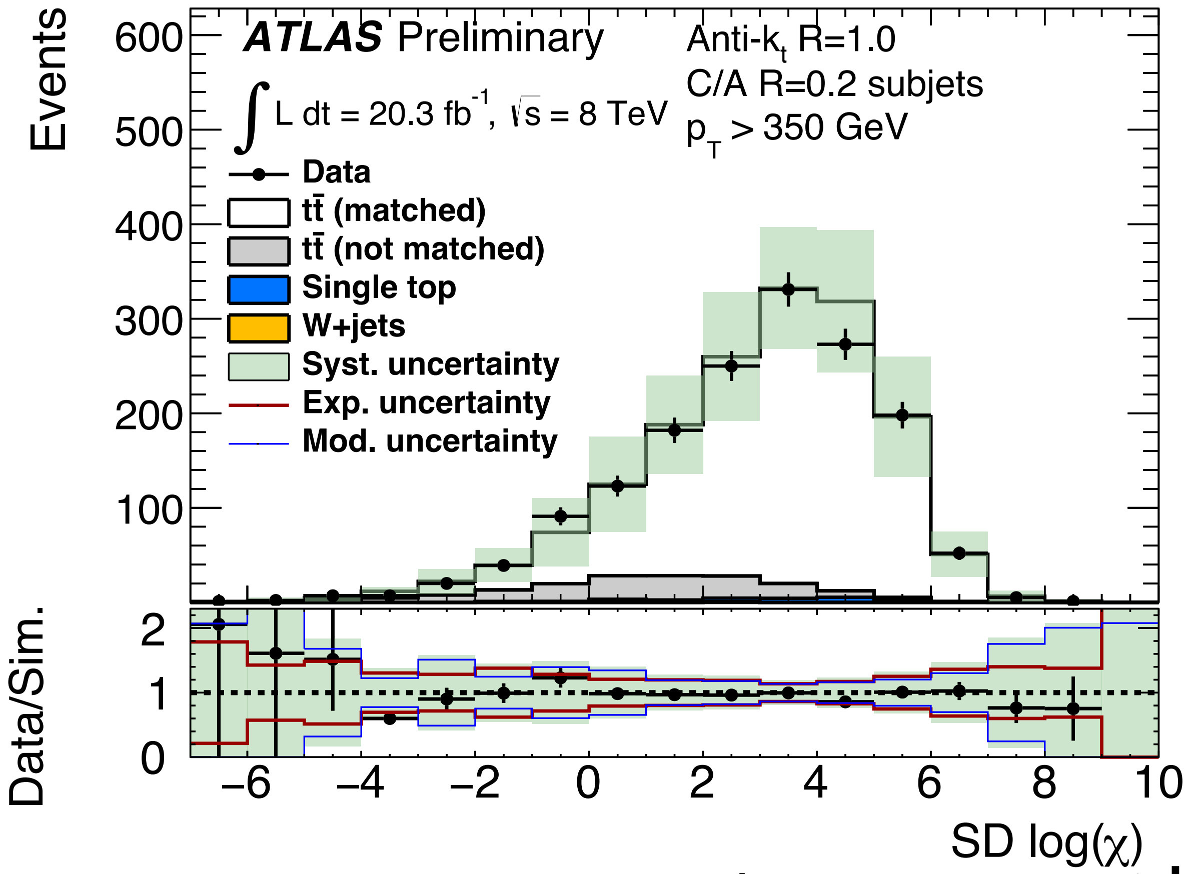


- Extensively used by 7-8 TeV CMS analyses
  - $t\bar{t}$ resonance search (arXiv:1506.03062)
  - High-p_T differential $t\bar{t}$ cross section measurement (CMS-PAS-TOP-14-012)

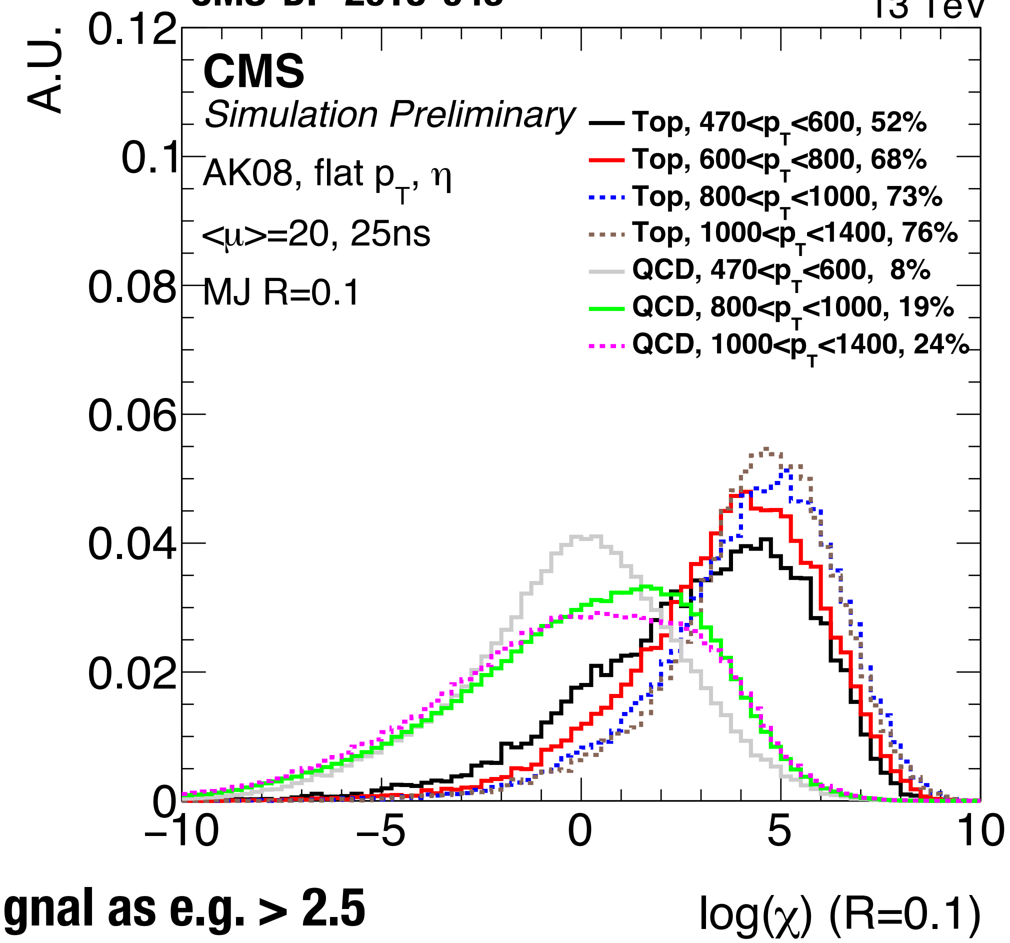
Shower Decomposition

- Likelihood that large-R jet is signal (hadronic top decay) vs background process (QCD)
  - Input is R=0.1-0.2 subjets
  - Assume each subjet comes from certain source of radiation (top decay, ISR, ...)
  - Calculate probability that subjet configuration comes from particular decay chain
- Discriminating variable: $\chi = \frac{\text{sum of signal probabilities}}{\text{sum of bkg probabilities}}$

ATLAS-CONF-2015-036



CMS-DP-2015-043



→ **tag signal as e.g. > 2.5**

N-Subjettiness

- Jet shape variable to measure consistency of jet to have N subjets

$$\tau_N = \frac{1}{d_0} \sum_k p_{\mathrm{T}k} \times \Delta R_k^{\min} \quad \text{with} \quad d_0 \equiv \sum_k p_{\mathrm{T}k} \times R$$

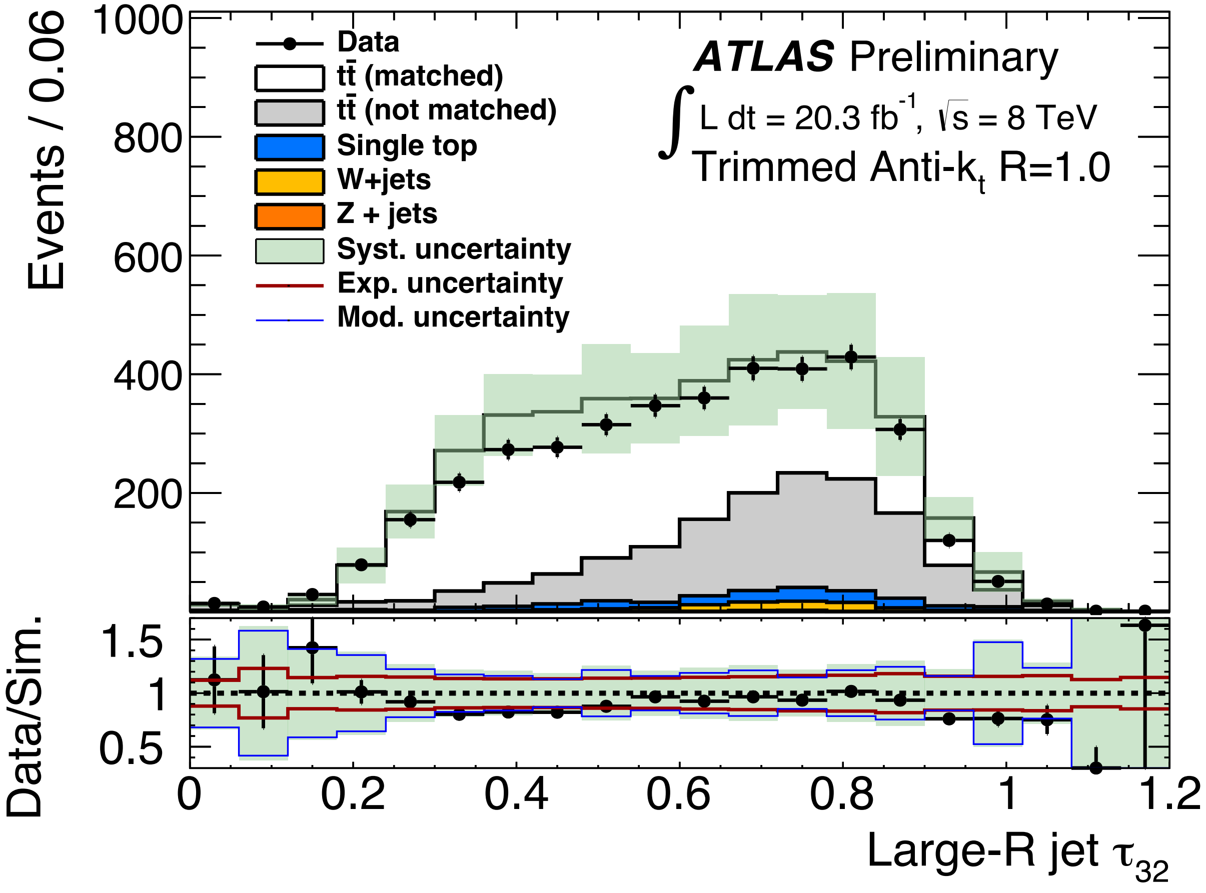
p_T of constituent k

distance between constituent k & axis of closest subjet

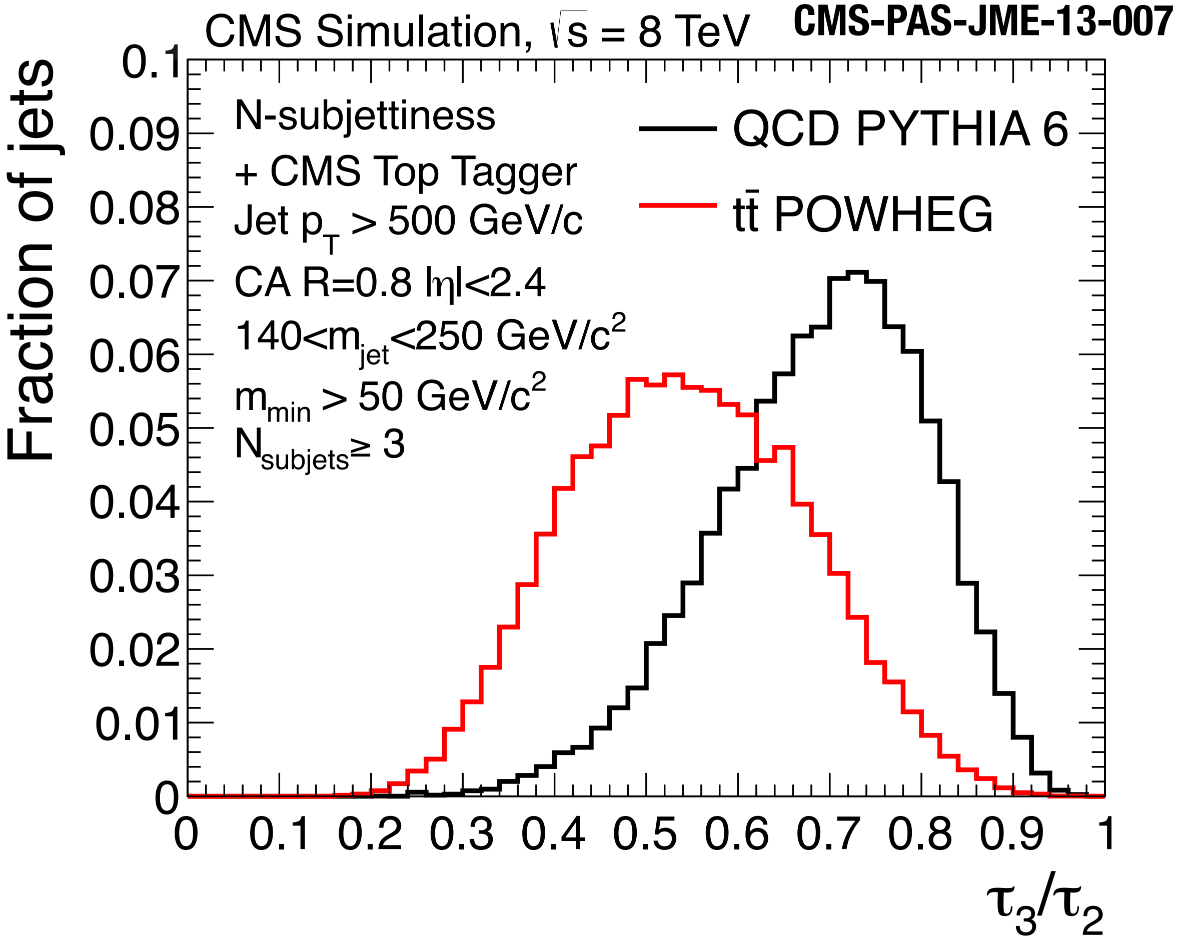
large-R jet distance parameter

- $\tau_{32} = \tau_3 / \tau_2$: discriminate 3-prong subjet structure (top jet) vs non-top jet
  - Typical cut value: $\tau_{32} < 0.6\text{-}0.7$

ATLAS-CONF-2015-036



CMS-PAS-JME-13-007



k_t Splitting Scale

- Measure of scale of last recombination step in k_t algorithm
  - k_t clusters high p_T & large-angle “proto jets” last

$$\sqrt{d_{ij}} = \min(p_{\mathrm{T}i}, p_{\mathrm{T}j}) \times \Delta R_{ij}$$

d_{12}: last step

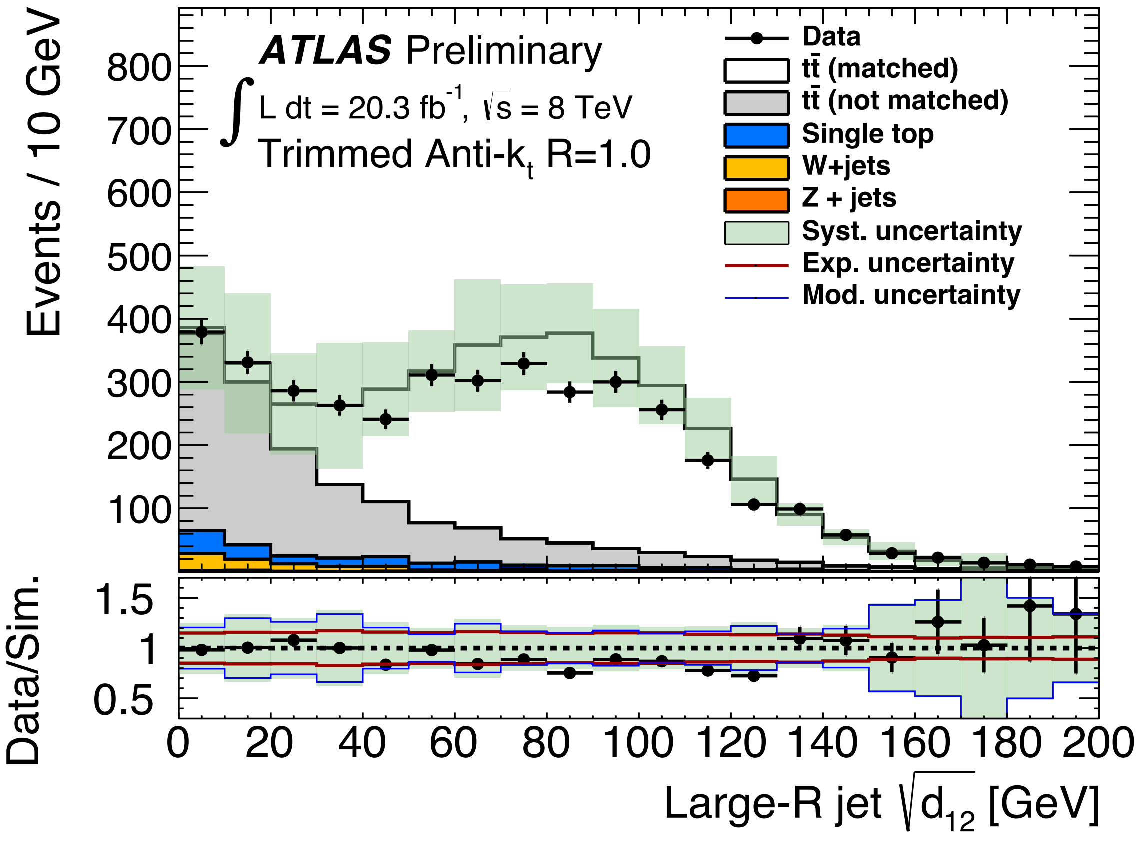
d_{23}: 2nd last step

subjet p_T

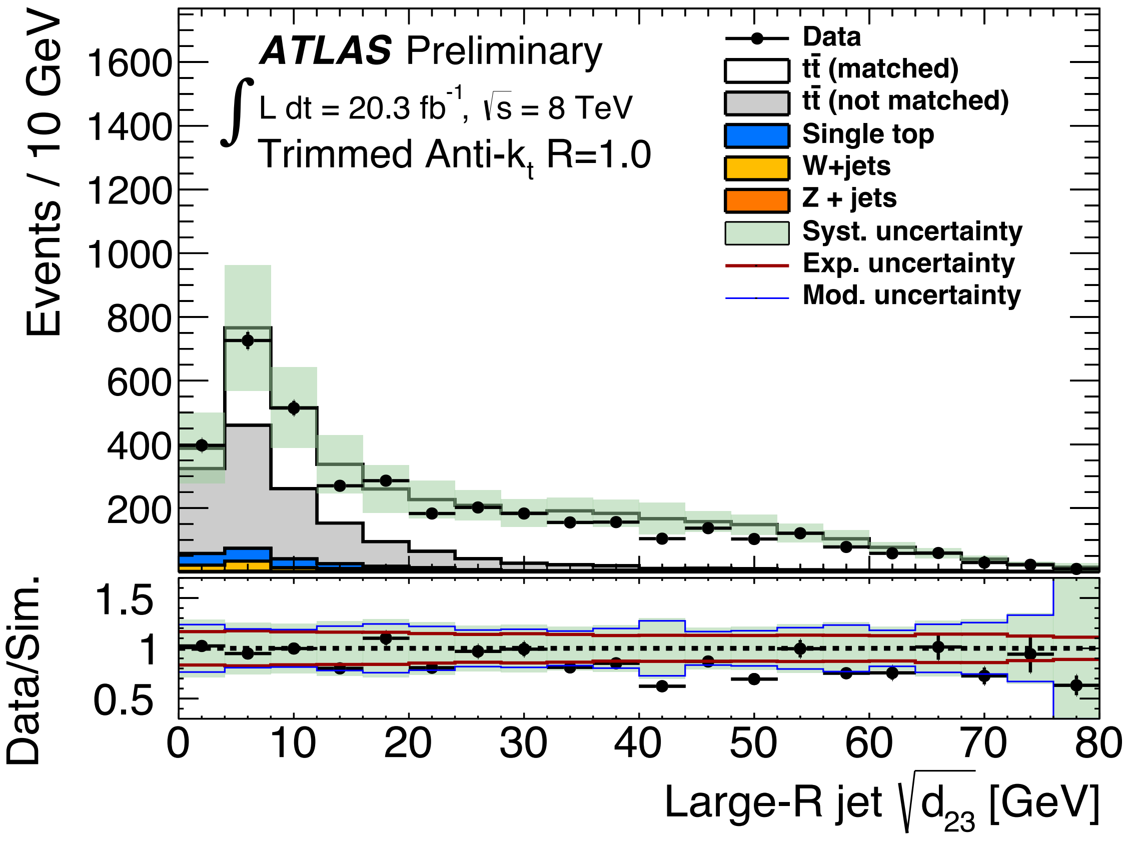
distance between 2 subjets

  - Typical cut value: $\sqrt{d_{12}} > 40$ GeV, $\sqrt{d_{23}} > 10$-20 GeV
  - Often used in ATLAS 8 TeV analyses

ATLAS-CONF-2015-036

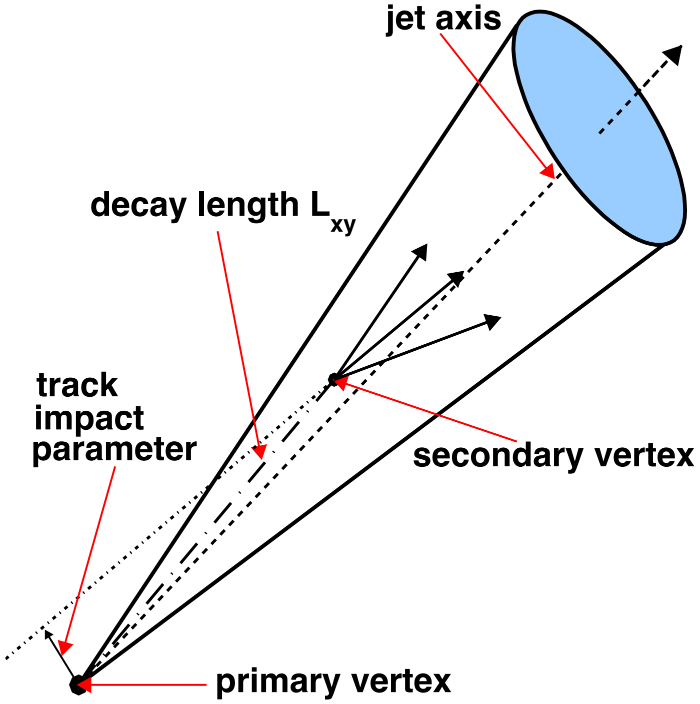


ATLAS-CONF-2015-036



Subjet b-Tagging

- Apply subjet b-tagging to increase QCD rejection
- Secondary vertex (SV) / track impact parameter (IP)

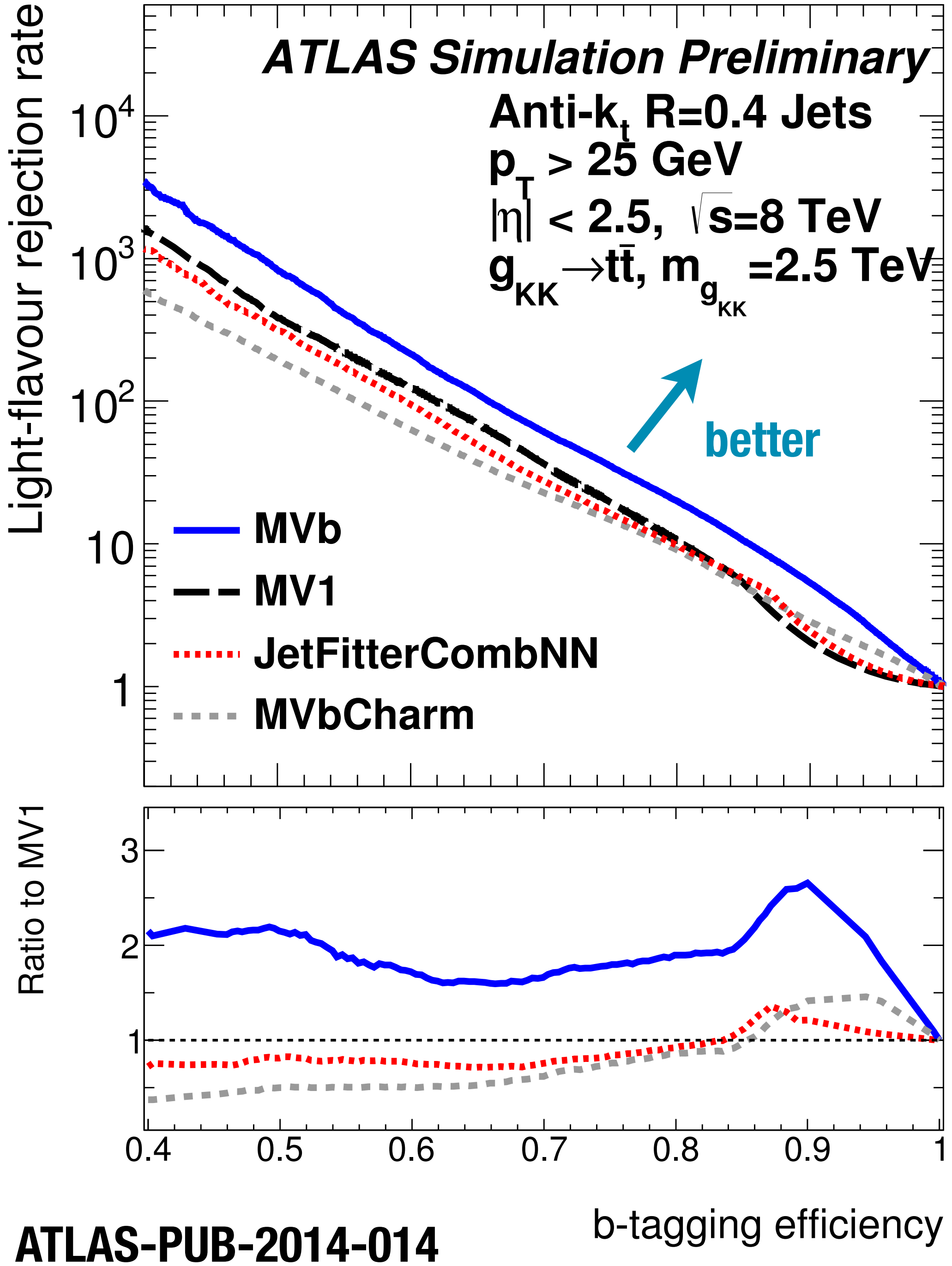


- **ATLAS**
  - Default: MV1 (neutral network based)
  - Improved tagger for additional discrimination in boosted regime: MVb
- **CMS**
  - Default: CSV (likelihood based), applied to either large-R jet vs each subjet
  - Improved tagger using all tracks to reconstruct SV, by construction independent of jet size: IVFCSV

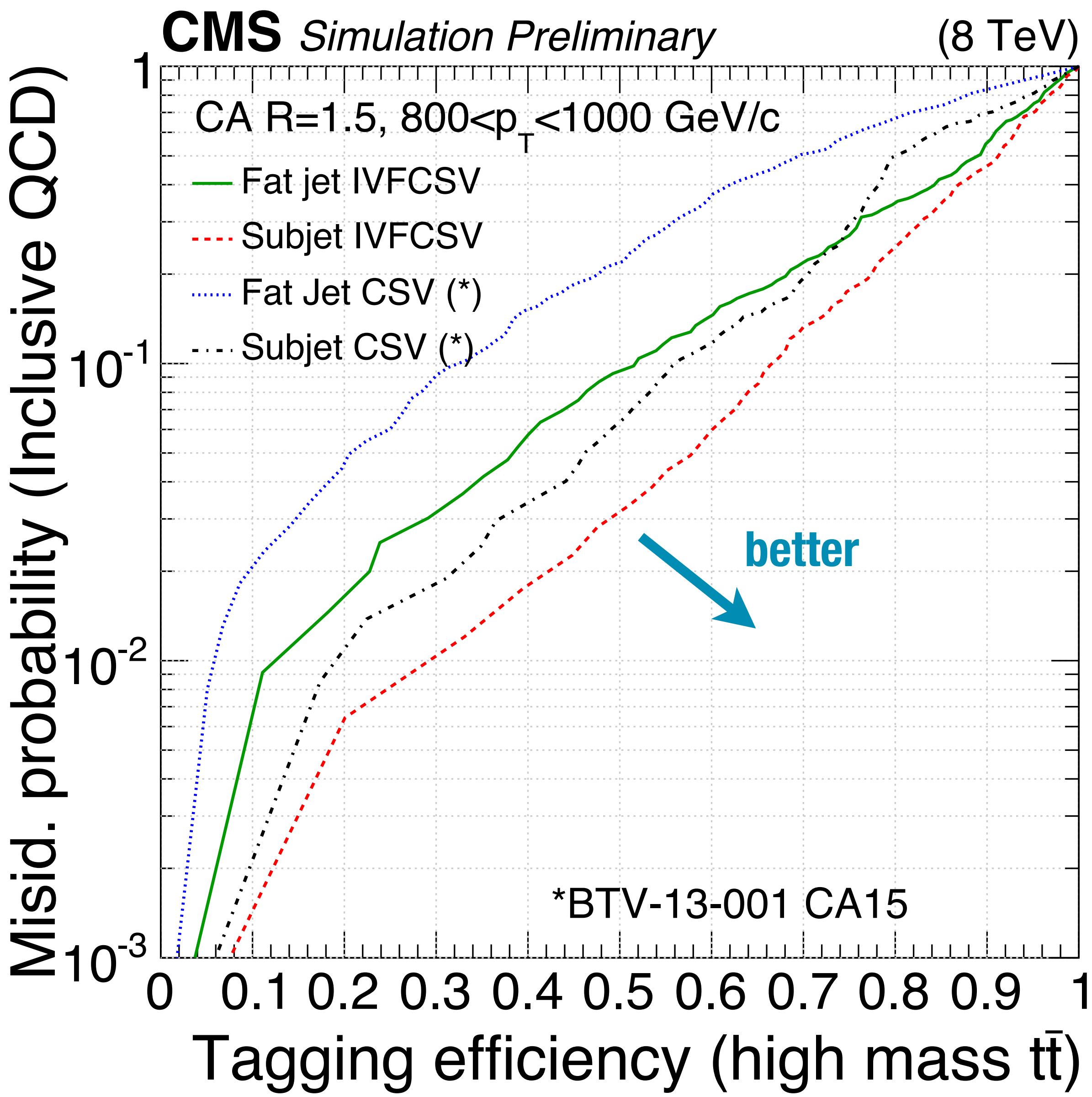
Subjet b-Tagging

ATLAS: improved tagger x2 better in boosted regime

CMS: better performance from updated tagger + applying b-tagging to subjet



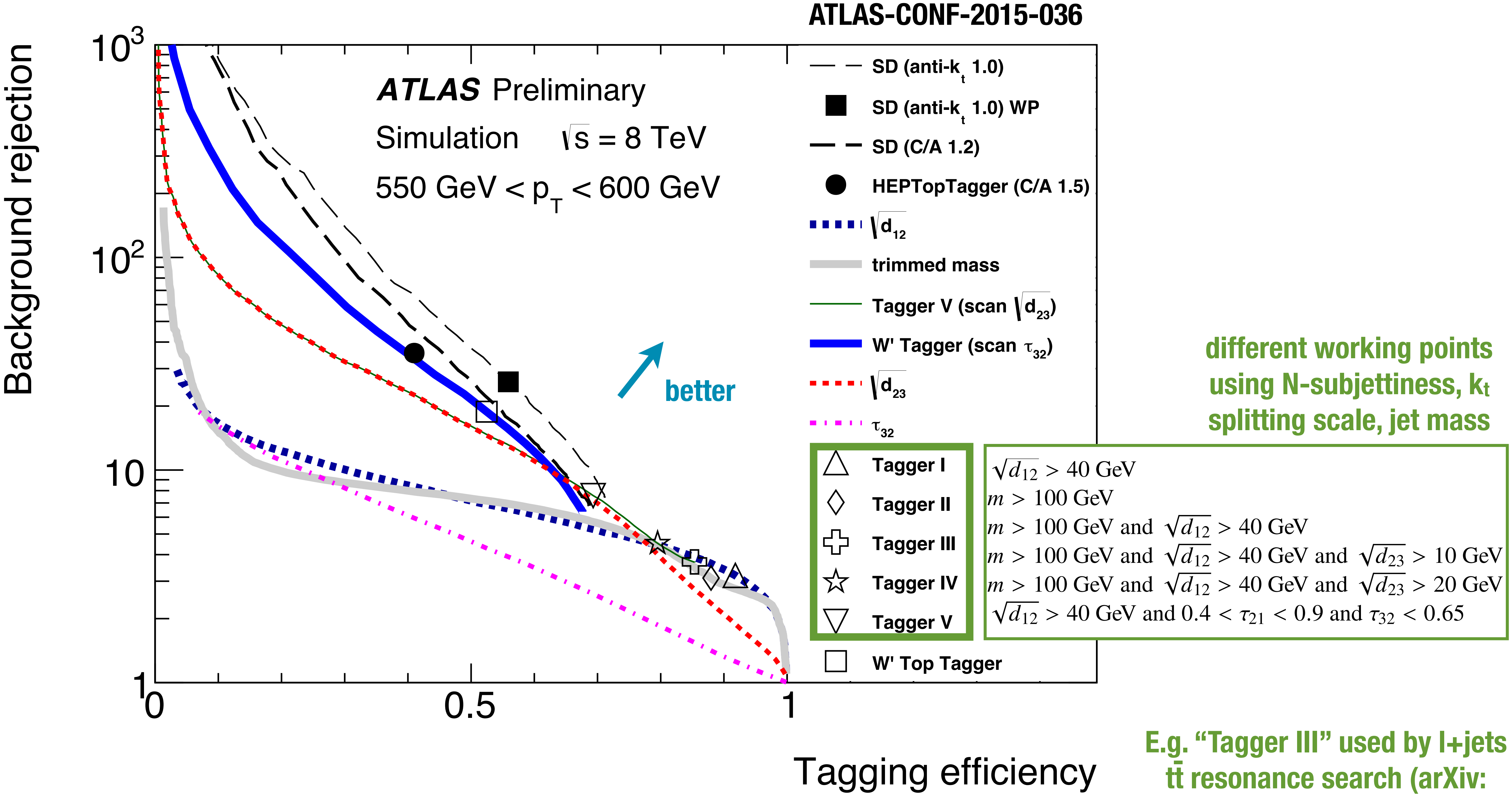
ATLAS-PUB-2014-014



CMS DP-2014/031

Performance Comparison (ATLAS)

ATLAS-CONF-2015-036



△ Tagger I	$\sqrt{d_{12}} > 40$ GeV
◇ Tagger II	$m > 100$ GeV
✚ Tagger III	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV
☆ Tagger IV	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 10$ GeV
▽ Tagger V	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 20$ GeV
	$\sqrt{d_{12}} > 40$ GeV and $0.4 < \tau_{21} < 0.9$ and $\tau_{32} < 0.65$

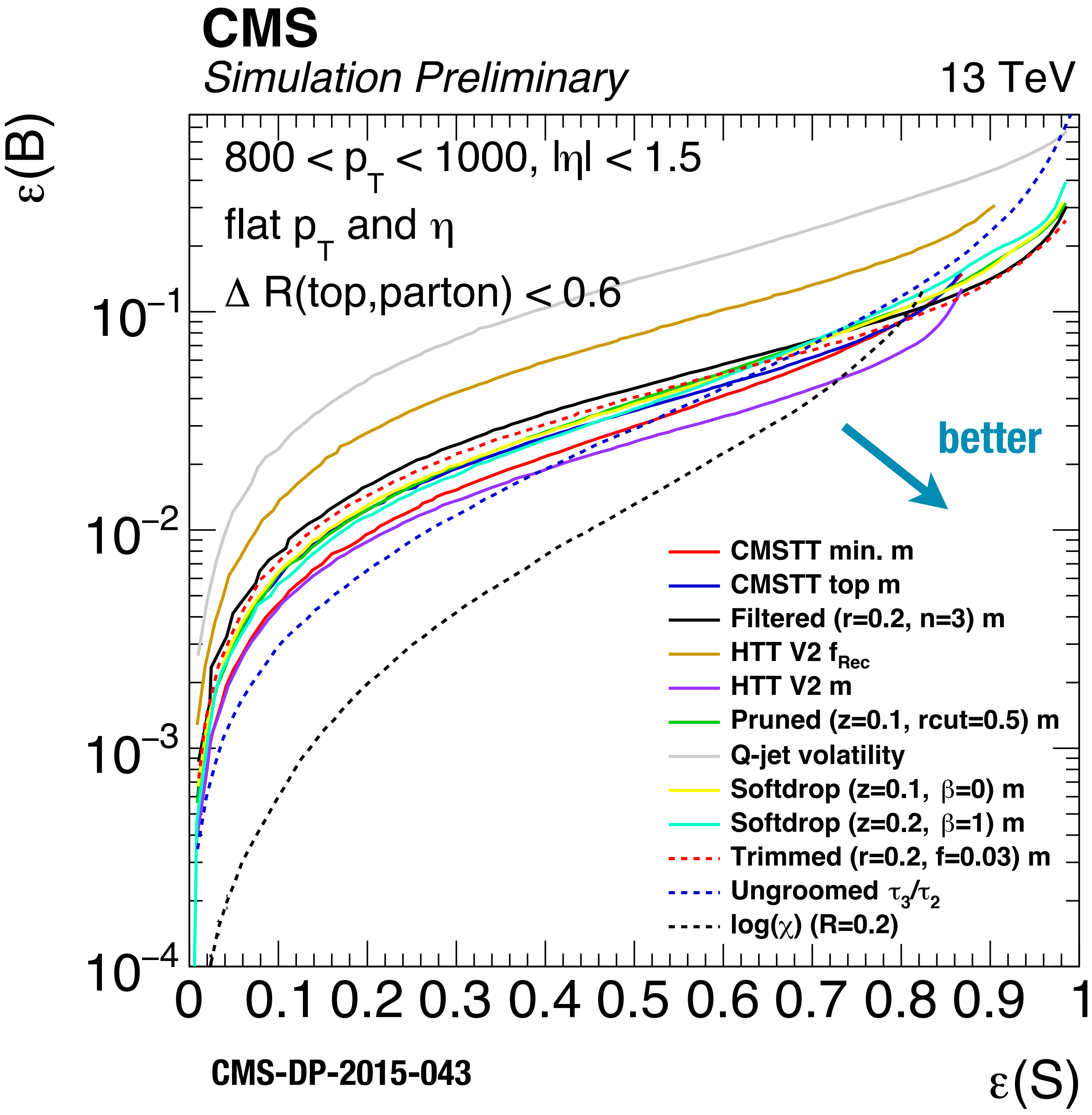
different working points using N-subjettiness, k_t splitting scale, jet mass

E.g. “Tagger III” used by l+jets $t\bar{t}$ resonance search (arXiv: 1505.07018) & high-p_T $t\bar{t}$ cross section (ATLAS-CONF-2014-057)

Performance Comparison (CMS)

Single Variables

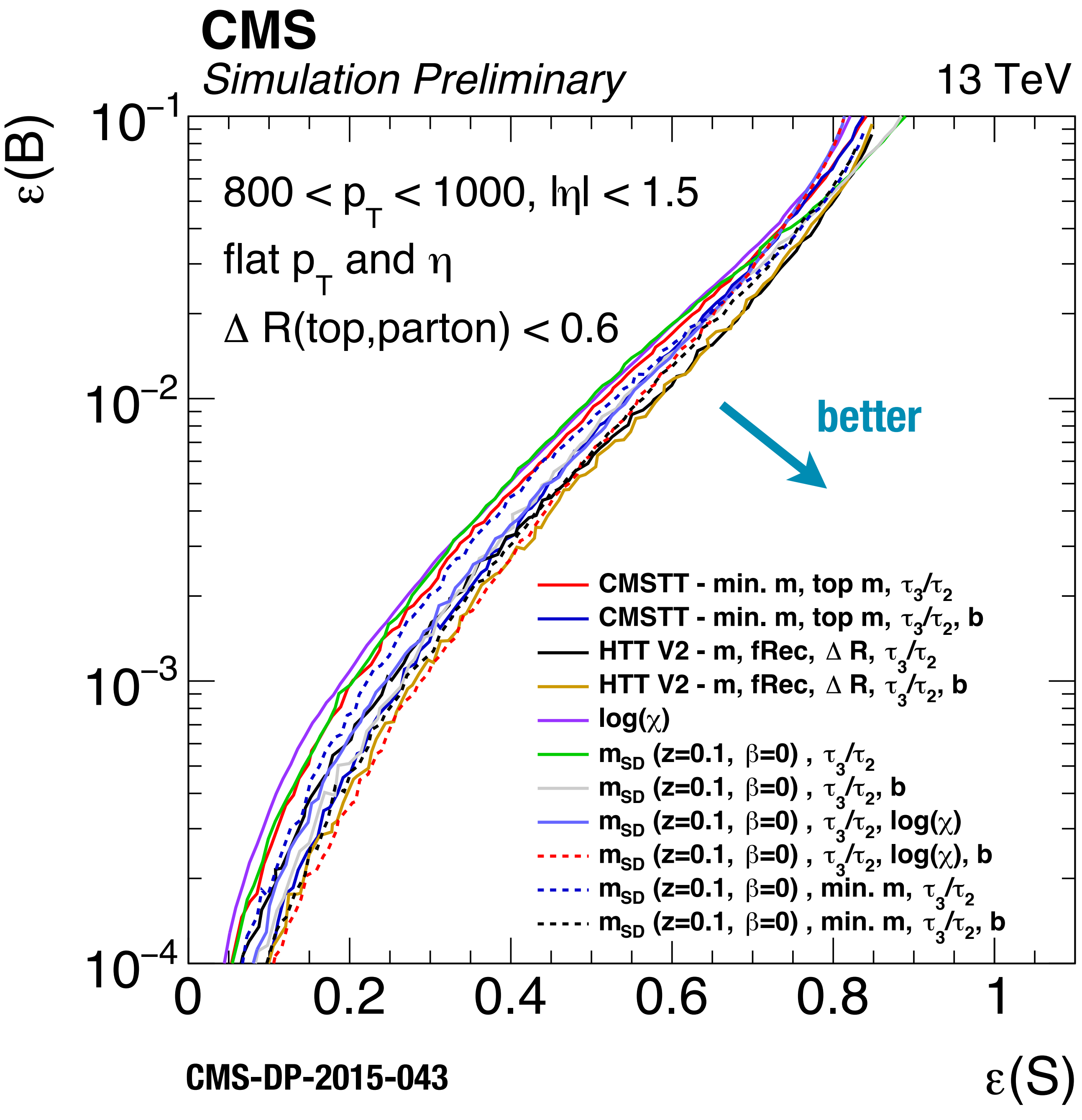
shower decomposition most discriminating variable



CMS-DP-2015-043

Combined Variables

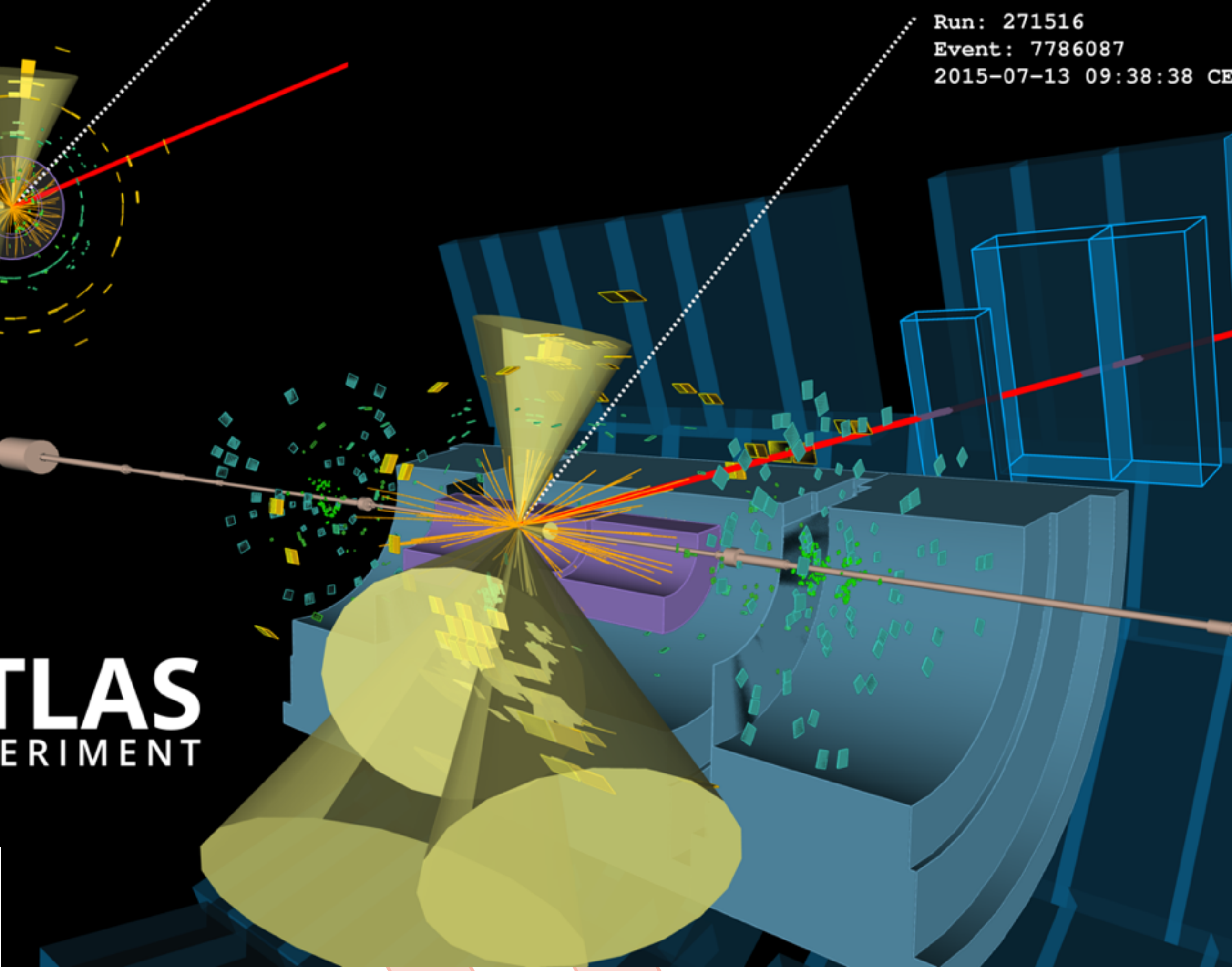
similar performance for variety of variables

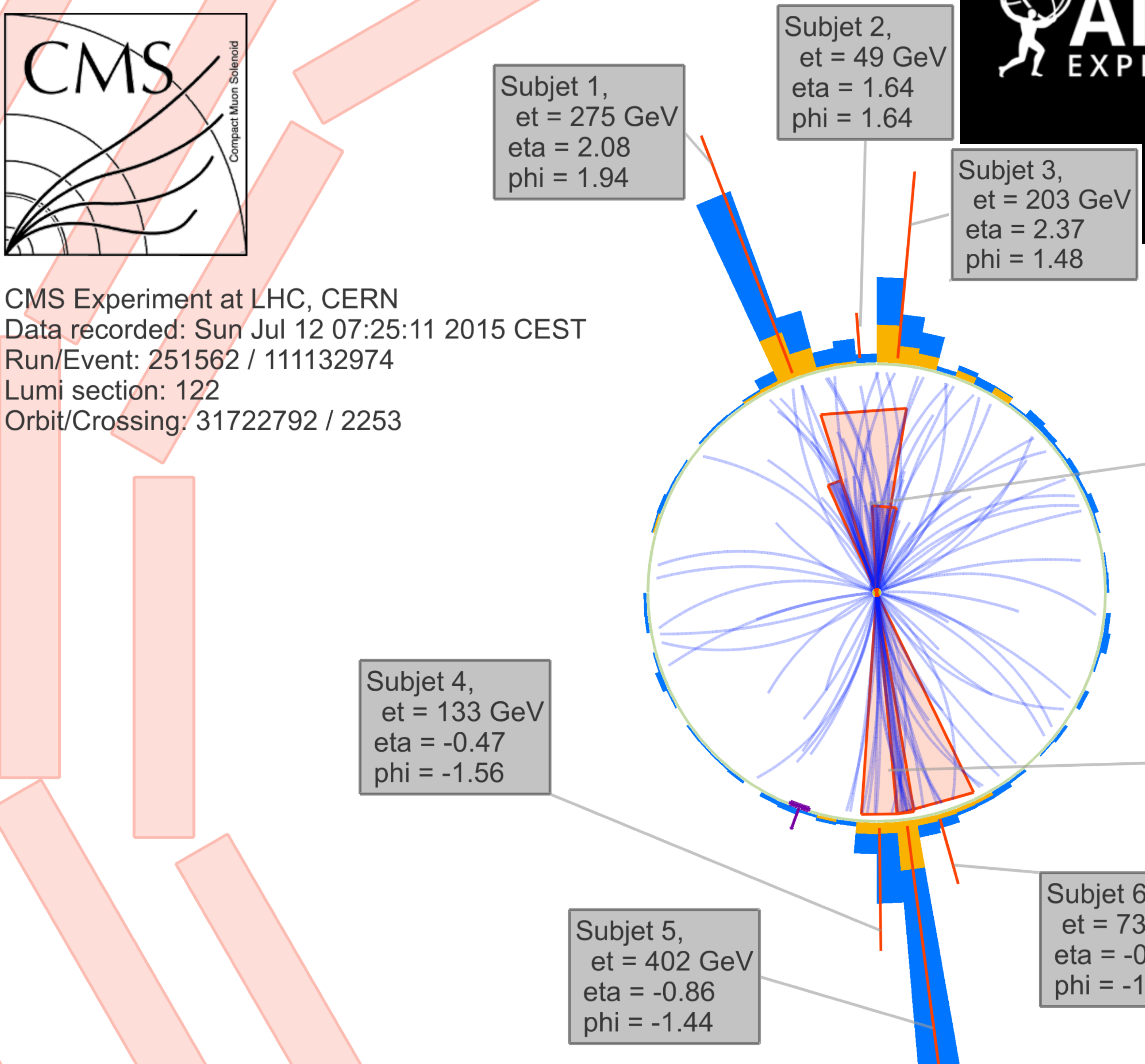


CMS-DP-2015-043

Conclusions

- Extensive list of tools!
- Substructure used in many BSM searches & (recently) measurements





CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253

Subjet 1,
et = 275 GeV
eta = 2.08
phi = 1.94

Subjet 2,
et = 49 GeV
eta = 1.64
phi = 1.64

Subjet 3,
et = 203 GeV
eta = 2.37
phi = 1.48

Top jet candidate 1
pt = 488 GeV
eta = 2.22
phi = 1.74
mass = 176 GeV

Top jet candidate 2,
pt = 613 GeV
eta = -0.70
phi = -1.46
mass = 177 GeV

Subjet 4,
et = 133 GeV
eta = -0.47
phi = -1.56

Subjet 5,
et = 402 GeV
eta = -0.86
phi = -1.44

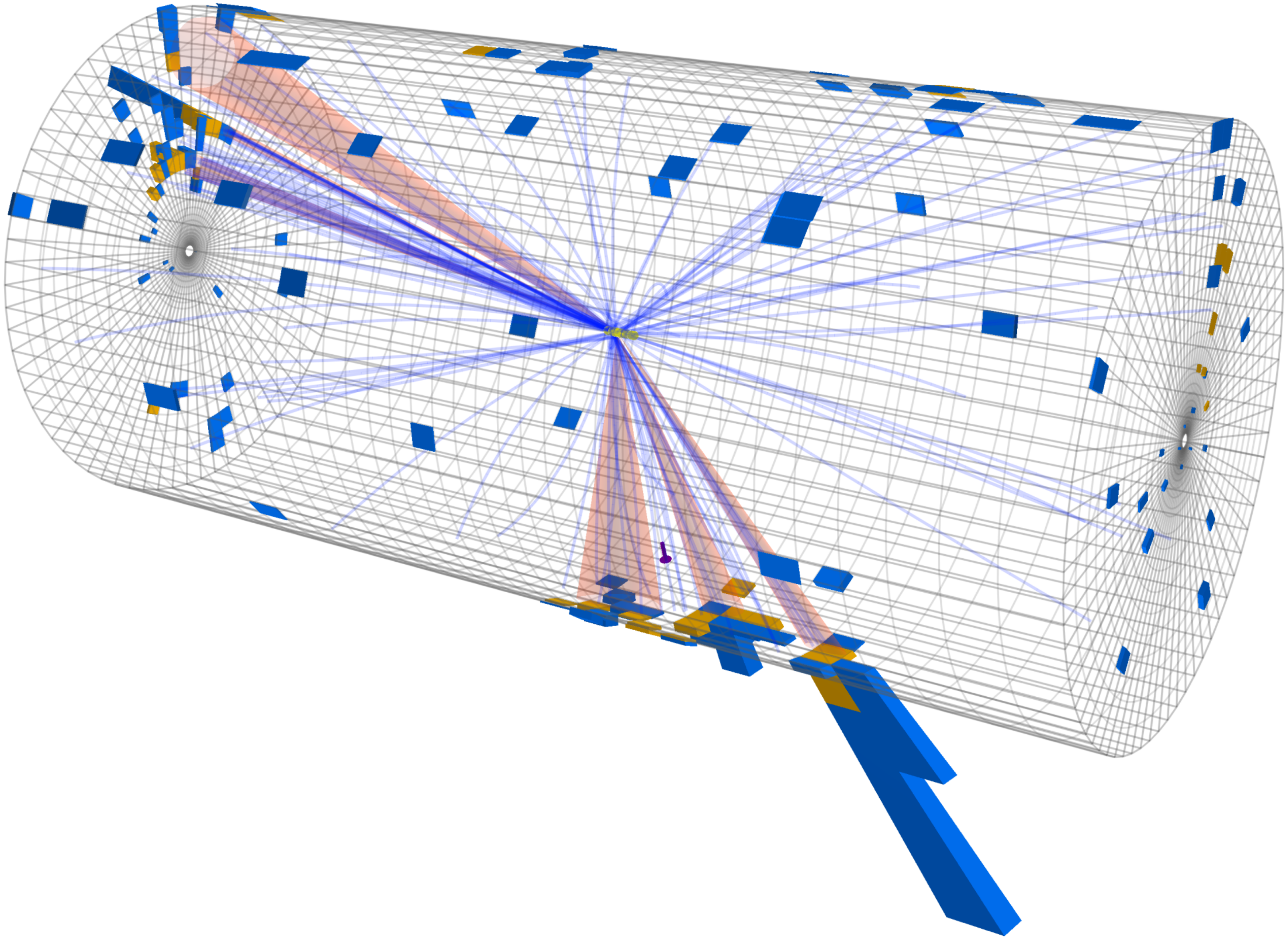
Subjet 6,
et = 73 GeV
eta = -0.18
phi = -1.30

- Important ingredient for 13 TeV physics!

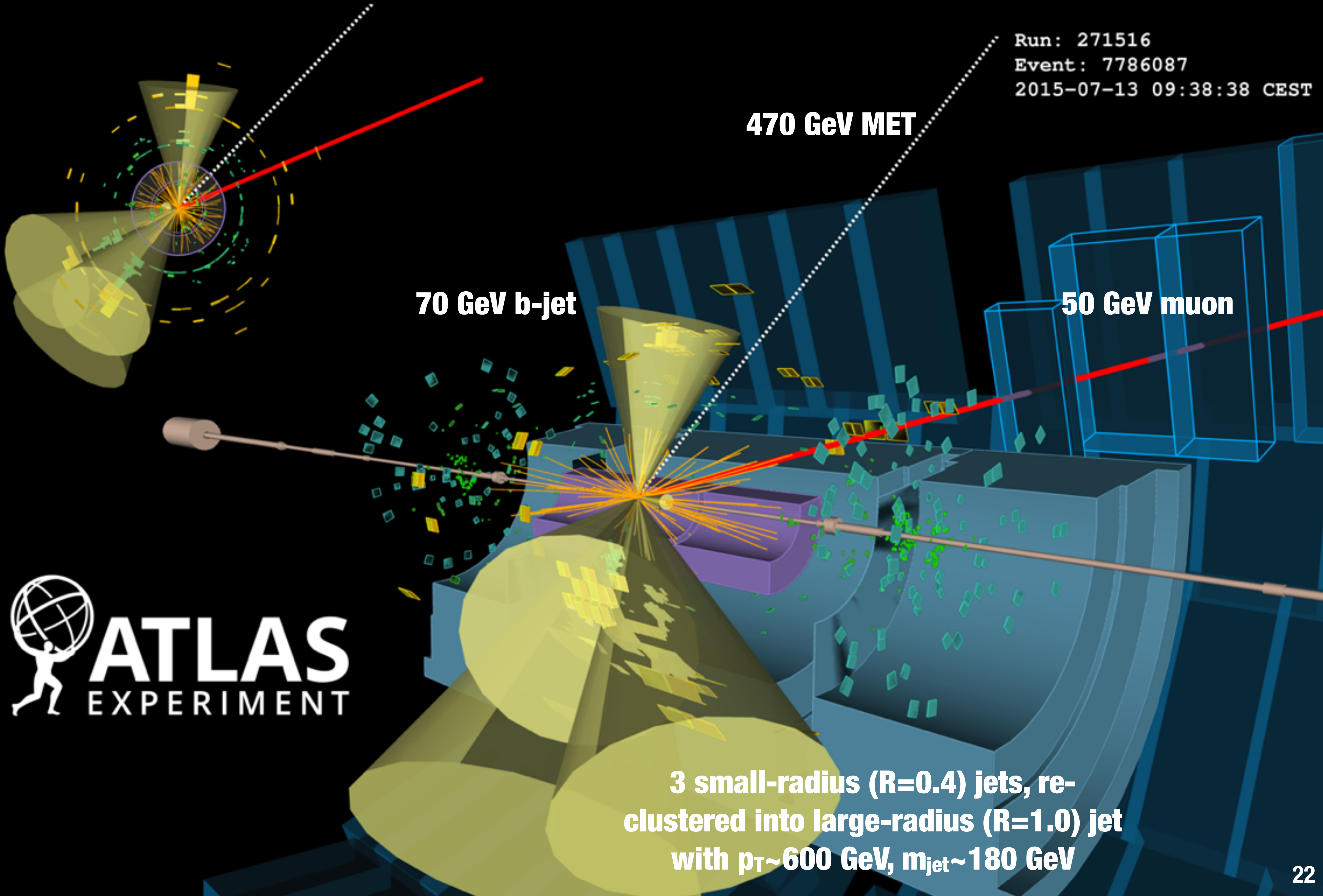
BACKUP

Run-2 Top-Jet Candidate (CMS)

CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253



Run-2 Top Jet Candidate (ATLAS)



Run: 271516
Event: 7786087
2015-07-13 09:38:38 CEST

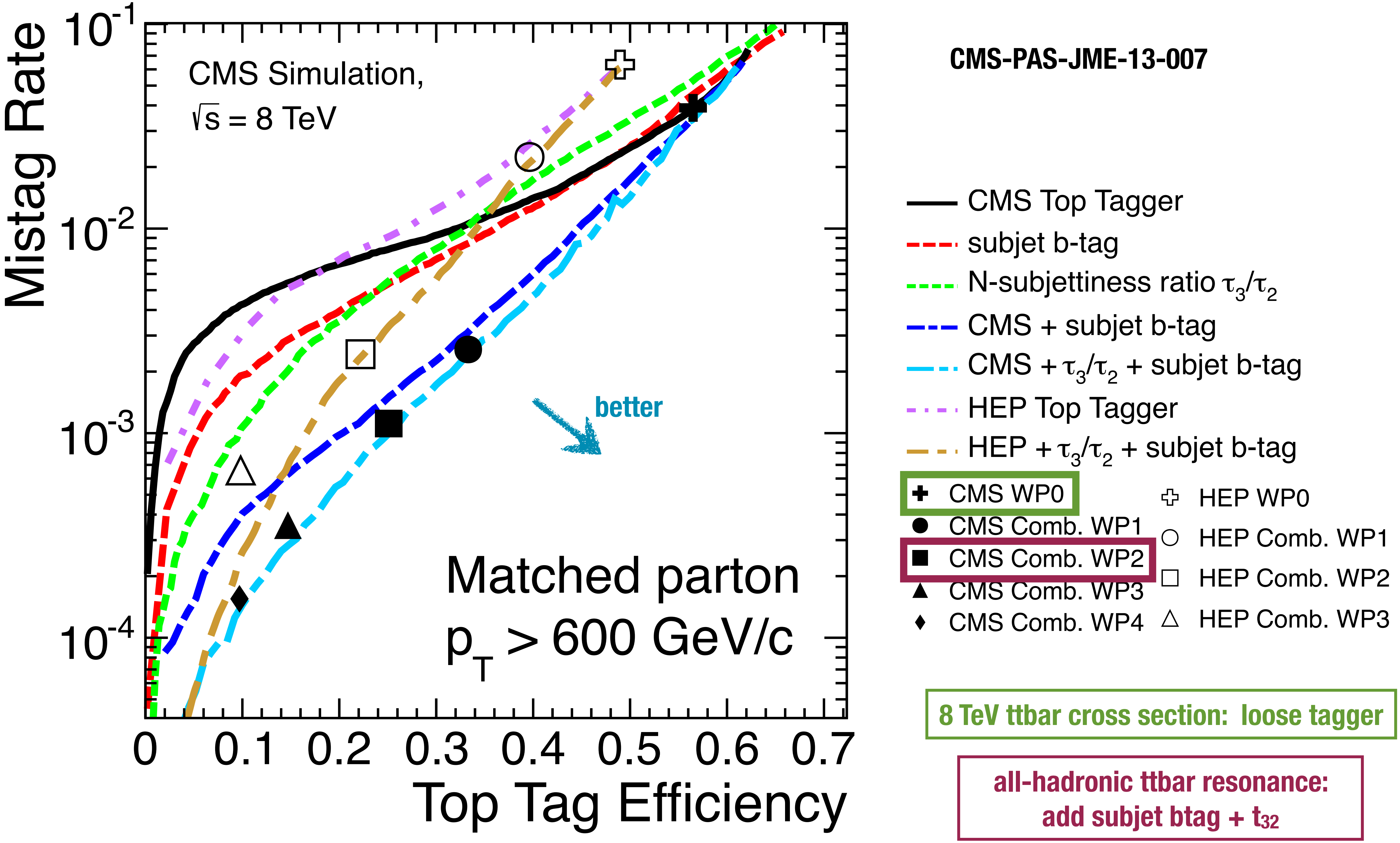
470 GeV MET

70 GeV b-jet

50 GeV muon

3 small-radius (R=0.4) jets, re-clustered into large-radius (R=1.0) jet with p_T~600 GeV, m_{jet}~180 GeV

Performance Comparison (CMS), 8 TeV



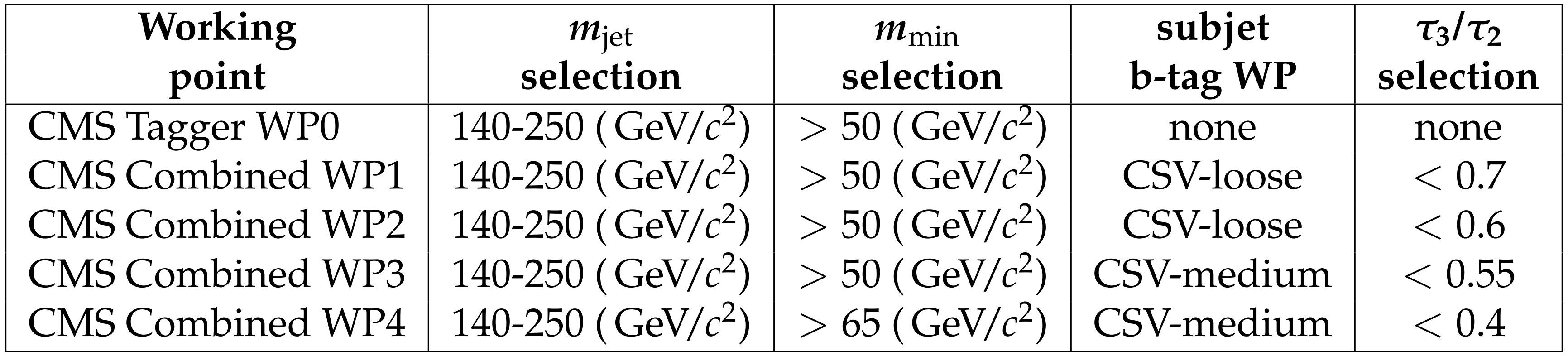
CMS-PAS-JME-13-007

8 TeV ttbar cross section: loose tagger

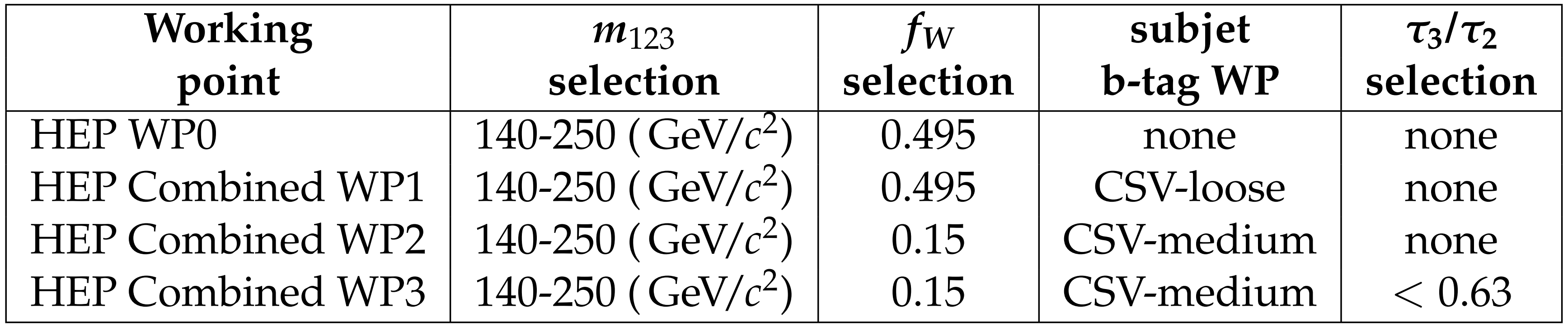
all-hadronic ttbar resonance:
add subjet btag + t_{32}

Top Tagger Working Points

CMS

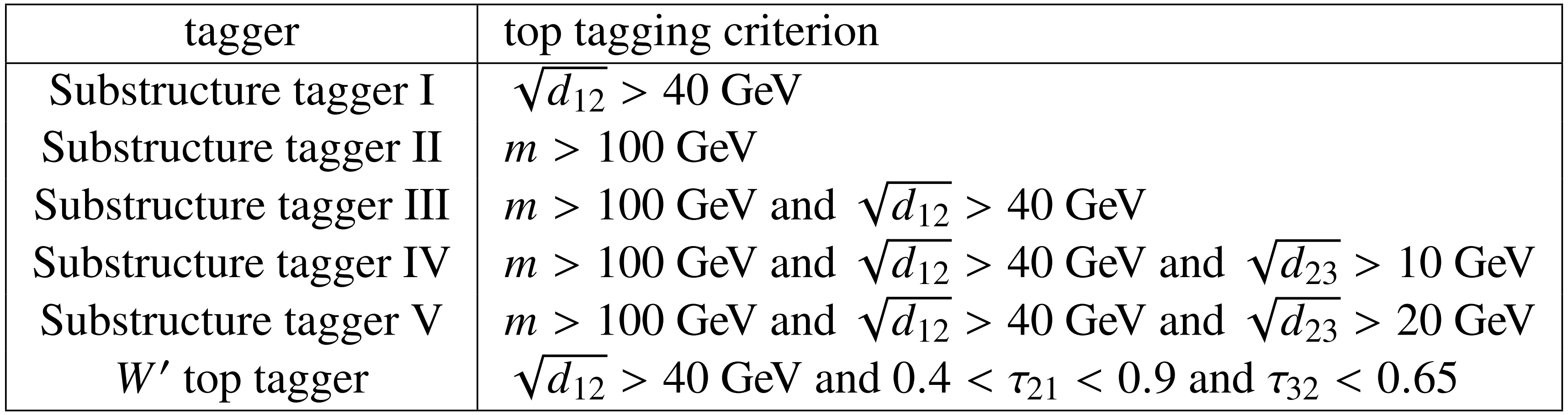


Working point	m_{jet} **selection**	m_{min} **selection**	**subjet b-tag WP**	τ_3/τ_2 **selection**
CMS Tagger WP0	140-250 (GeV/c^2)	$>$ 50 (GeV/c^2)	none	none
CMS Combined WP1	140-250 (GeV/c^2)	$>$ 50 (GeV/c^2)	CSV-loose	$<$ 0.7
CMS Combined WP2	140-250 (GeV/c^2)	$>$ 50 (GeV/c^2)	CSV-loose	$<$ 0.6
CMS Combined WP3	140-250 (GeV/c^2)	$>$ 50 (GeV/c^2)	CSV-medium	$<$ 0.55
CMS Combined WP4	140-250 (GeV/c^2)	$>$ 65 (GeV/c^2)	CSV-medium	$<$ 0.4



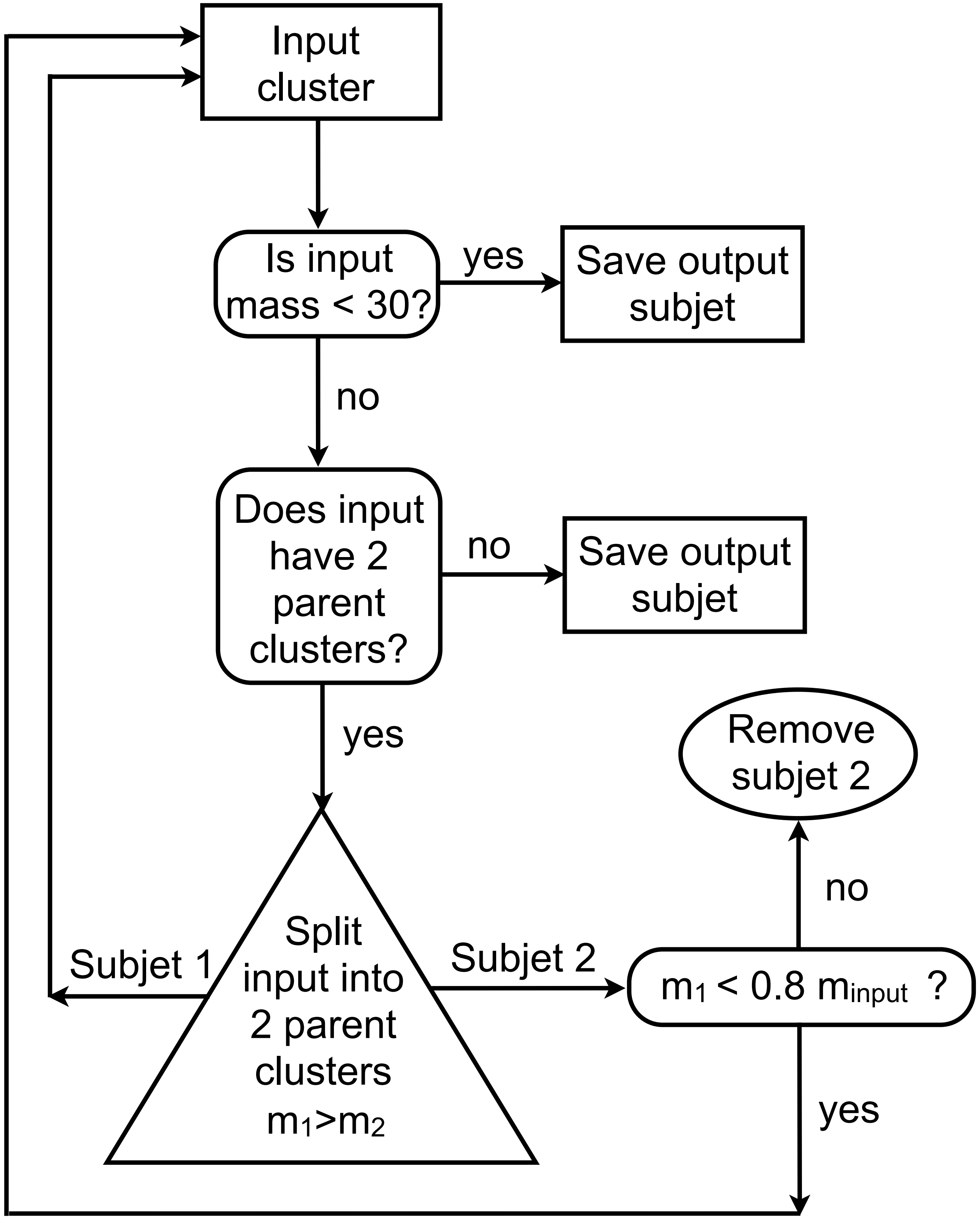
Working point	m_{123} **selection**	f_W **selection**	**subjet b-tag WP**	τ_3/τ_2 **selection**
HEP WP0	140-250 (GeV/c^2)	0.495	none	none
HEP Combined WP1	140-250 (GeV/c^2)	0.495	CSV-loose	none
HEP Combined WP2	140-250 (GeV/c^2)	0.15	CSV-medium	none
HEP Combined WP3	140-250 (GeV/c^2)	0.15	CSV-medium	$<$ 0.63

ATLAS



tagger	top tagging criterion
Substructure tagger I	$\sqrt{d_{12}} > 40$ GeV
Substructure tagger II	$m > 100$ GeV
Substructure tagger III	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV
Substructure tagger IV	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 10$ GeV
Substructure tagger V	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 20$ GeV
W' top tagger	$\sqrt{d_{12}} > 40$ GeV and $0.4 < \tau_{21} < 0.9$ and $\tau_{32} < 0.65$

HEP Top Tagger: Mass Drop Decomposition



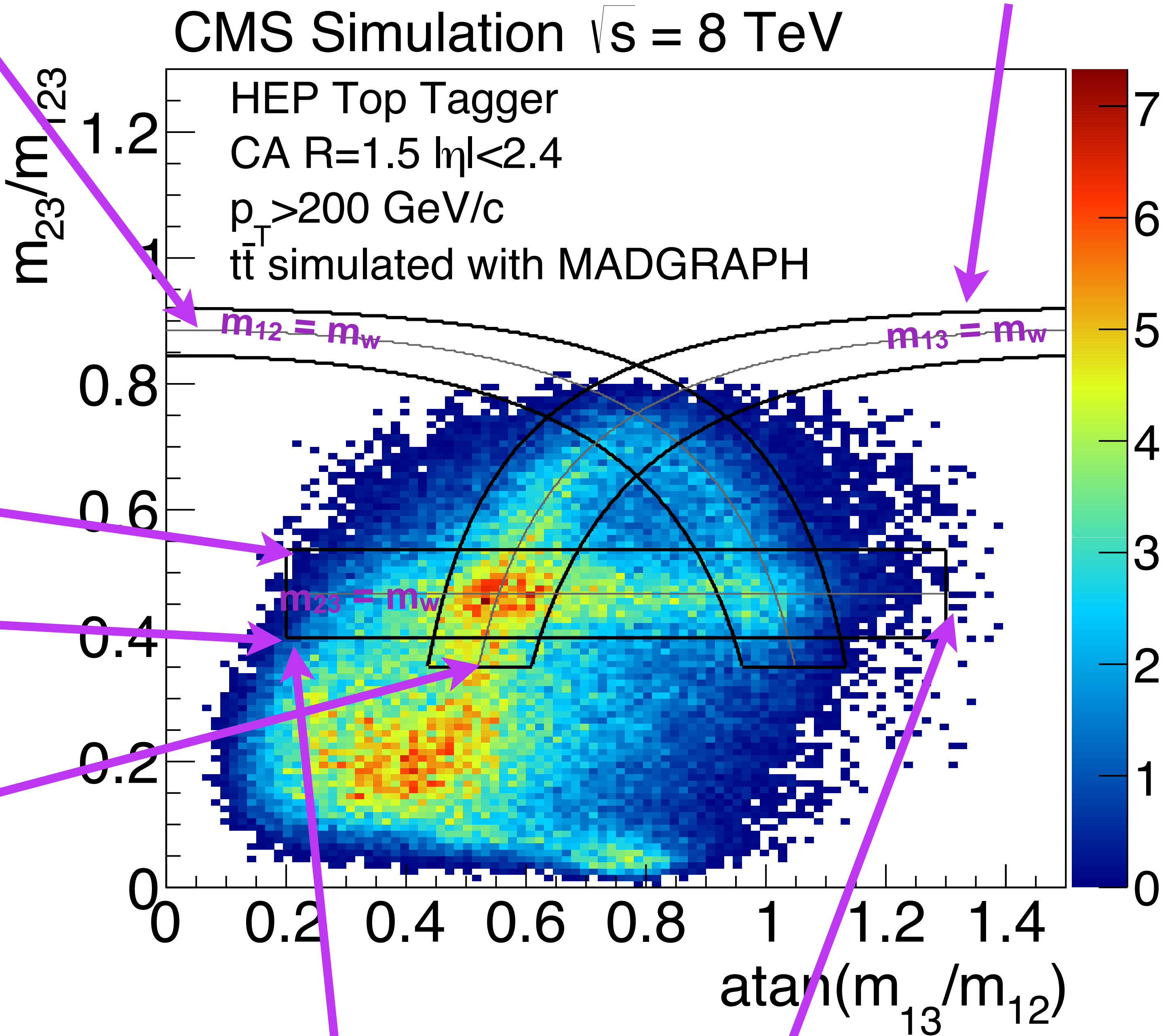
HEP Top Tagger: W Mass Selection

CMS-PAS-JME-13-007

Bi-dimensional distribution based on ratio of subjet pairwise masses

$$R^2_{\min}(1+(\frac{m_{12}}{m_{13}})^2) < 1-(\frac{m_{23}}{m_{123}})^2 < R^2_{\max}(1+(\frac{m_{12}}{m_{13}})^2)$$

$$R^2_{\min}(1+(\frac{m_{13}}{m_{12}})^2) < 1-(\frac{m_{23}}{m_{123}})^2 < R^2_{\max}(1+(\frac{m_{13}}{m_{12}})^2)$$



$$R_{\min} < \frac{m_{23}}{m_{123}} < R_{\max}$$

$$R_{\max} = (1+f_W) \times m_W/m_t$$

$$R_{\min} = (1-f_W) \times m_W/m_t$$

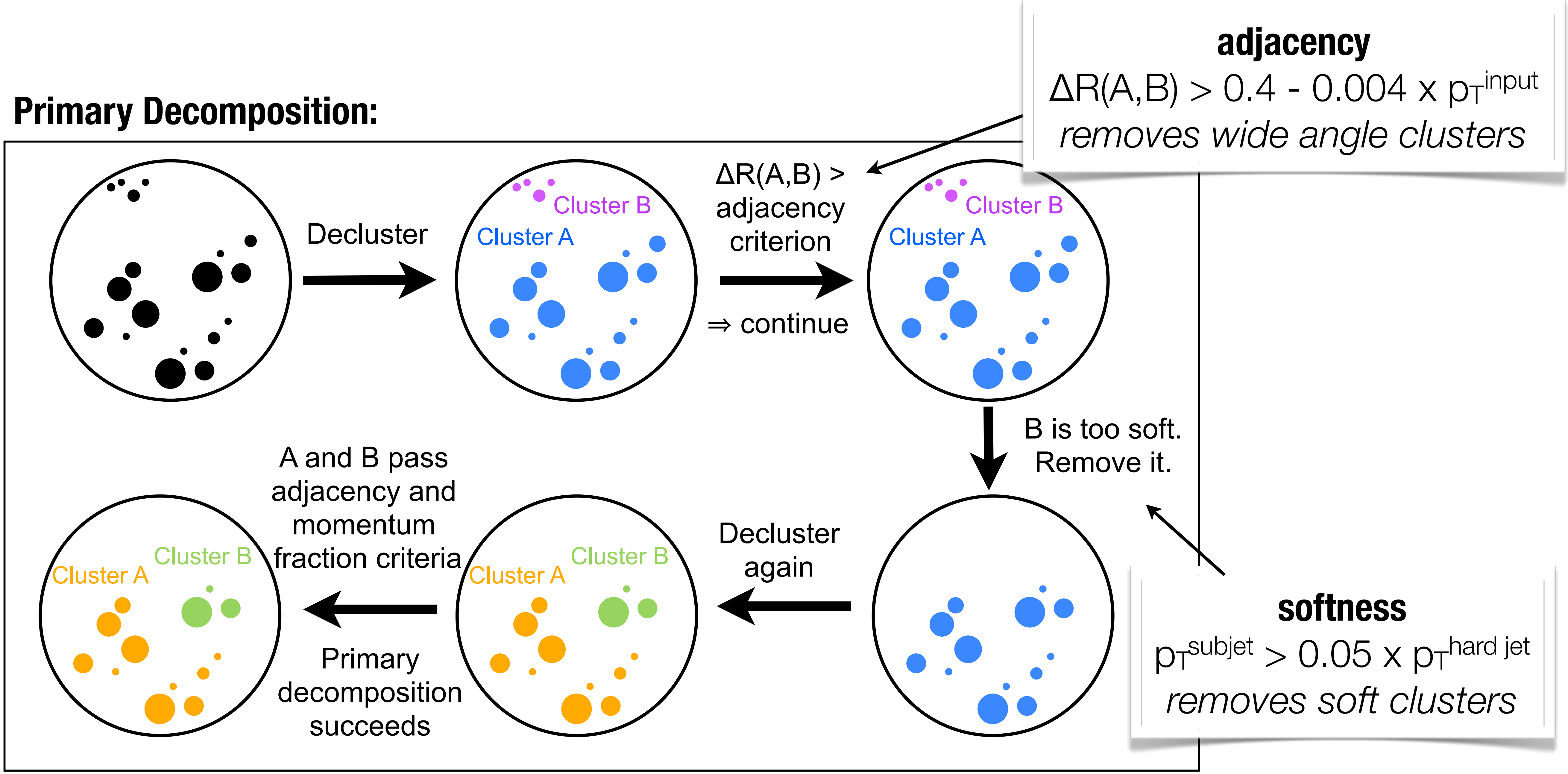
$$\frac{m_{23}}{m_{123}} > 0.35$$

$$0.2 < \arctan\frac{m_{13}}{m_{12}} < 1.3$$

CMS Top Tagger

CMS-PAS-JME-13-007

- Decomposition of jets into up to 4 subjets
- Input R=0.8 Cambridge-Aachen jets (used for $p_T > 400$ GeV)



Primary Decomposition:

adjacency

$\Delta R(A,B) > 0.4 - 0.004 \times p_T^{input}$

removes wide angle clusters

softness

$p_T^{subjet} > 0.05 \times p_T^{hard\ jet}$

removes soft clusters

Secondary Decomposition: Individually decluster A/B for 3 final subjets