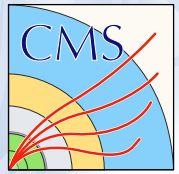
Boosted top: experimental tools overview

Emanuele Usai
on behalf of the **CMS** and **ATLAS** collaborations
prepared with the help of **Johannes Erdmann**

Universität Hamburg

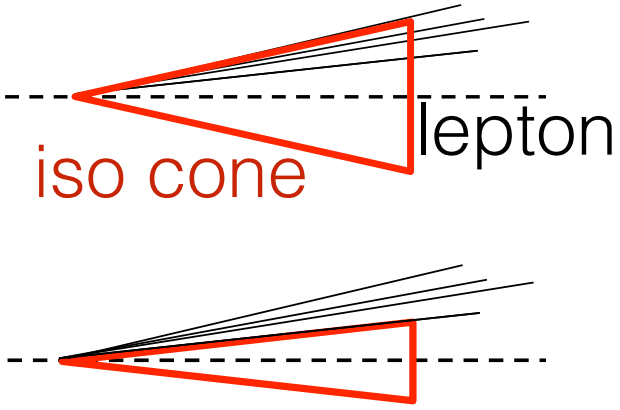


2014

7th International Workshop on Top Quark Physics

Cannes, October 1st, 2014

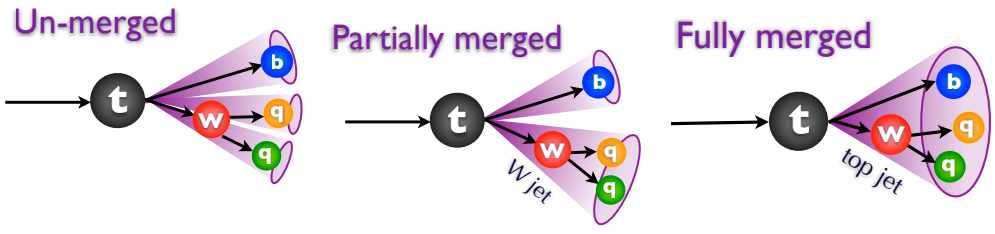
Strategy

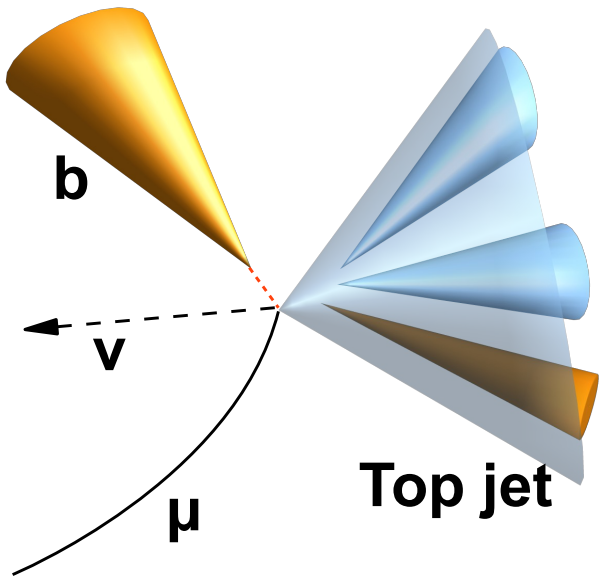


Boosted leptonic top

- lepton and b-jet very close
- classic jet isolation not optimal
- shrink the cone depending on the p_T of the lepton

Boosted hadronic top $\rightarrow$ main topic of the talk





- cluster whole decay in a large radius jet
- use top taggers
- add jet substructure variables on top of it

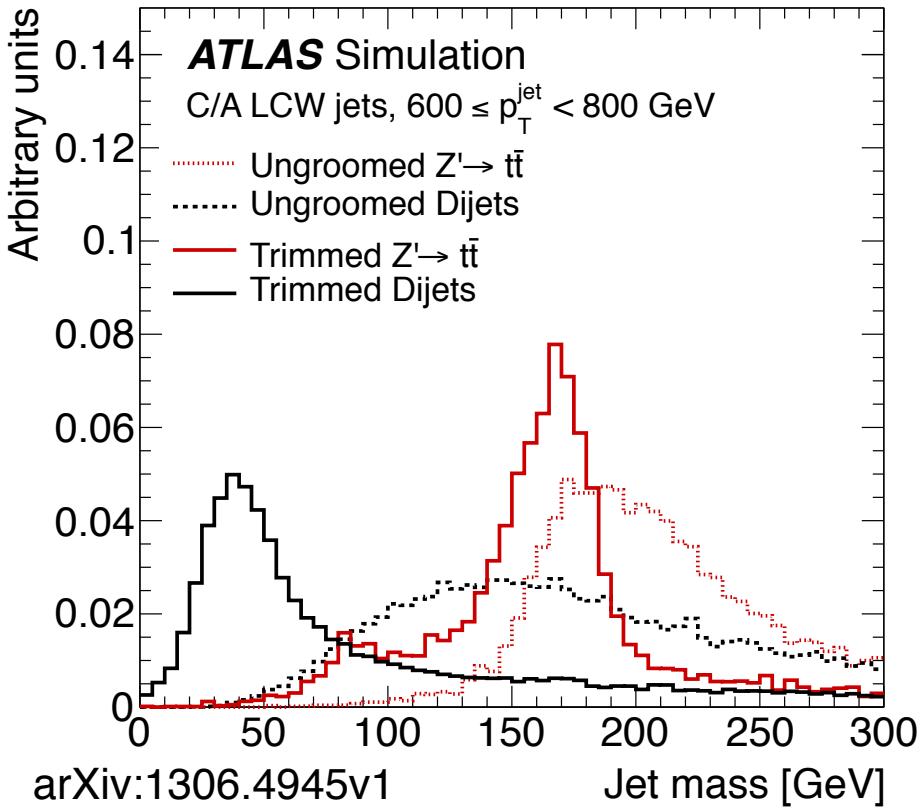
Jet grooming

Large radius jets:

- collect lots of soft QCD radiation
- suppress it to resolve hard decay product
- use jet grooming algorithms to remove soft, wide angle radiation
- mass drop filtering, pruning, trimming

Trimming:

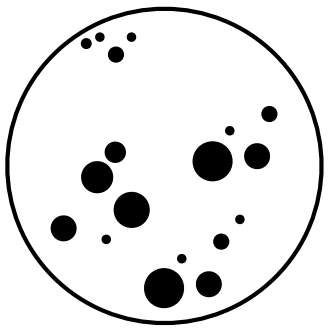
- k_t algorithm to cluster subjet of radius R_{sub}
- reject softer subjets using a p_T fraction requirement ($p_{T,i}/p_{T,jet} < f$)



Top taggers

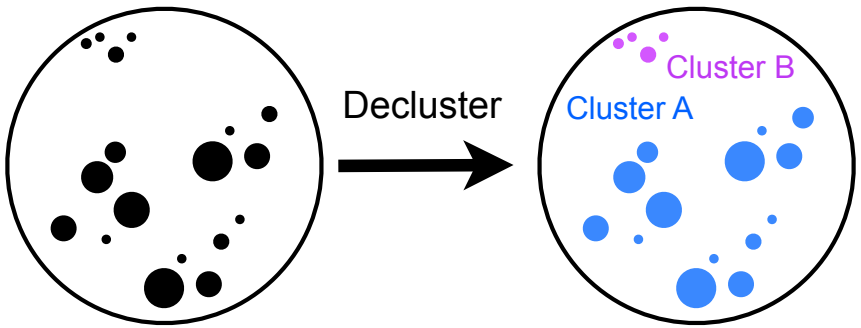
CMS Top Tagger (Kaplan et al., arXiv:0806.0848)

- start from Cambridge-Aachen R=0.8 jet



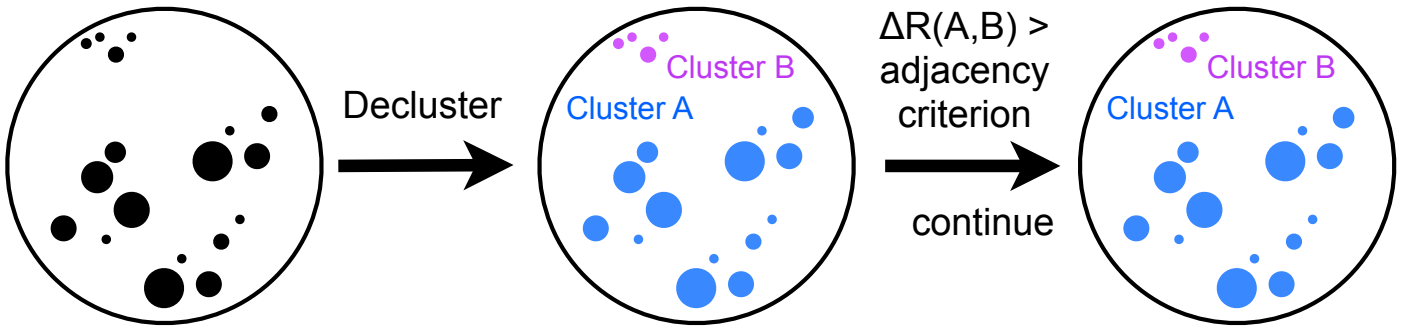
CMS Top Tagger (Kaplan et al., arXiv:0806.0848)

- start from Cambridge-Aachen R=0.8 jet



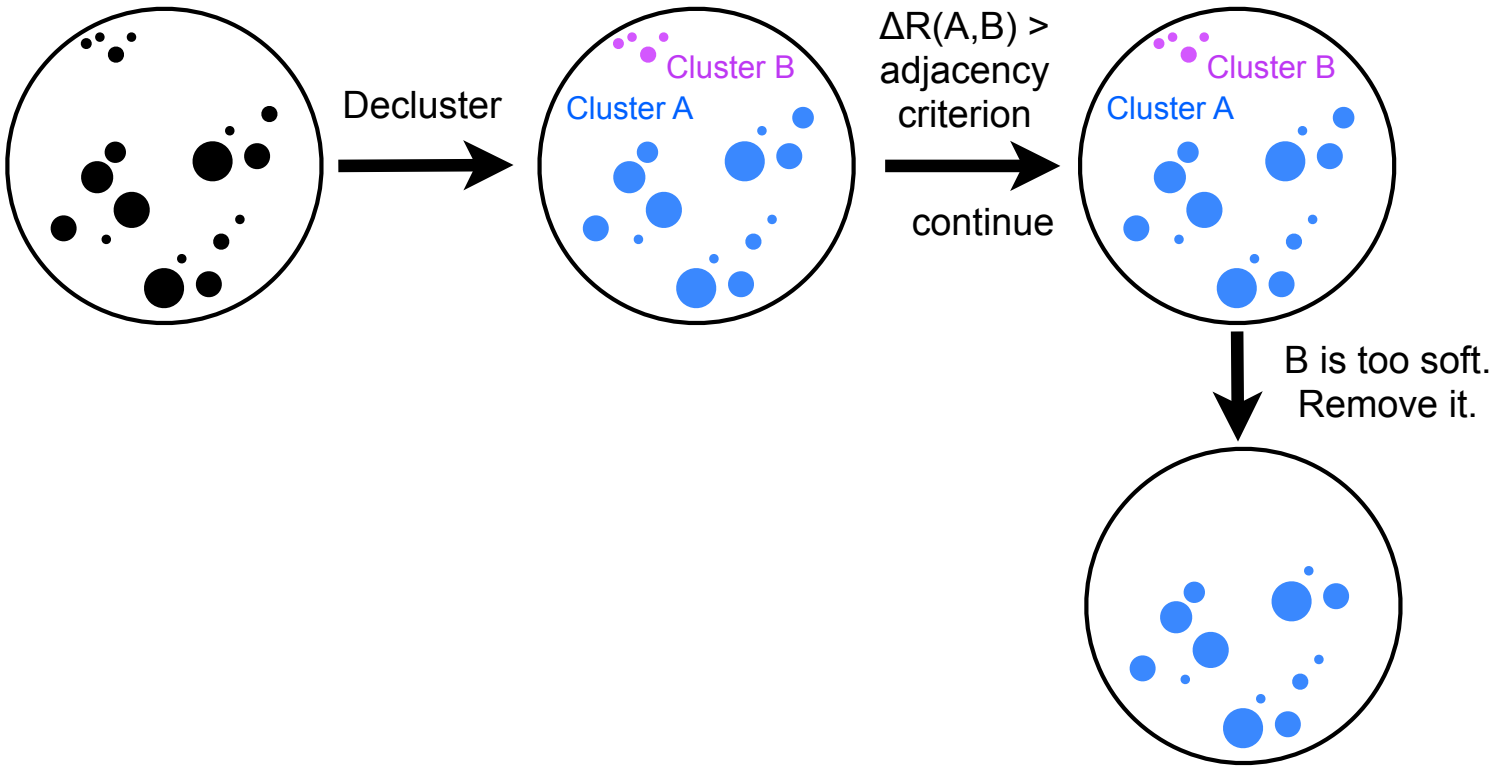
CMS Top Tagger (Kaplan et al., arXiv:0806.0848)

- start from Cambridge-Aachen R=0.8 jet
- adjacency: $\Delta R(A,B) > 0.4 - 0.004 \times p_T^{\text{input}}$, do not decluster if fails



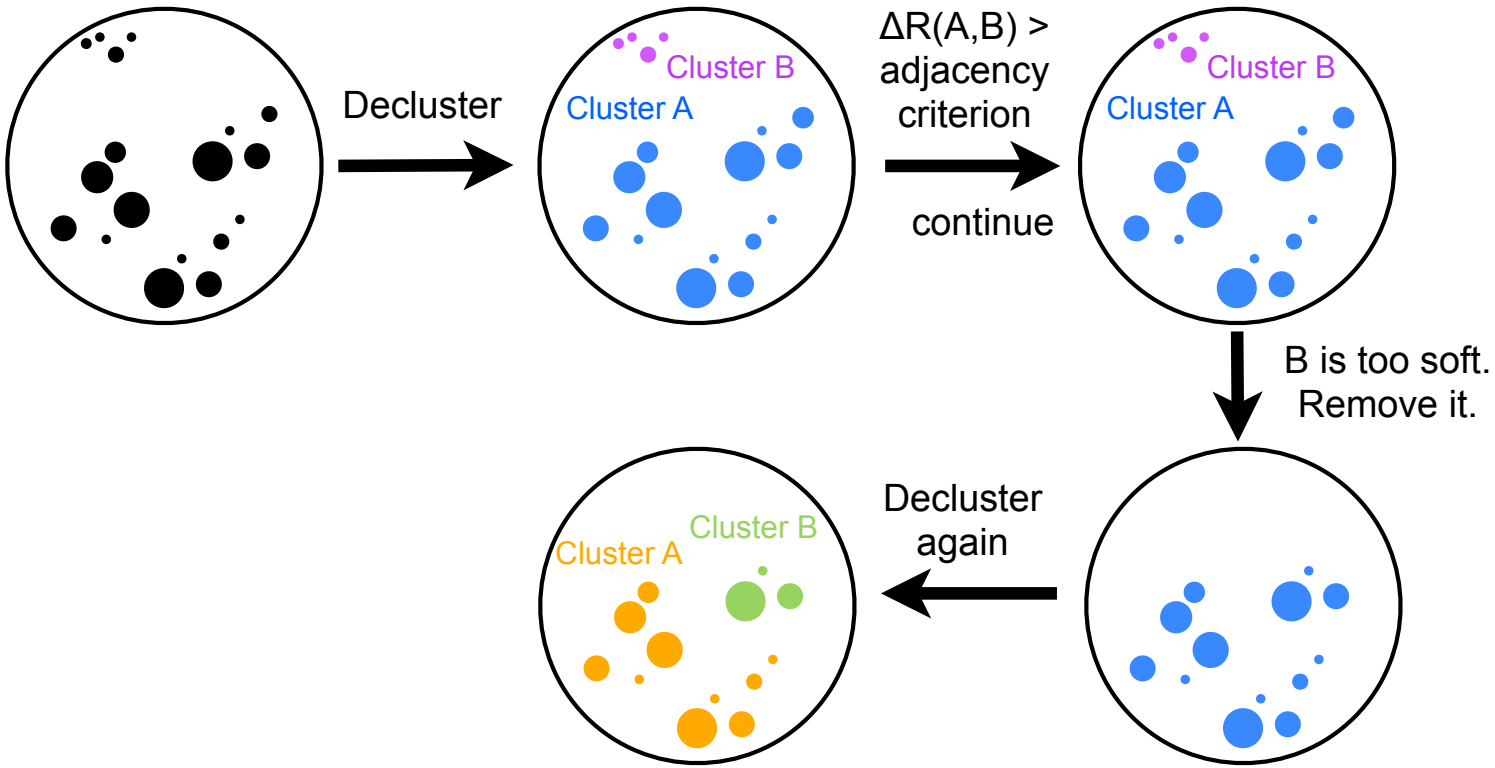
CMS Top Tagger (Kaplan et al., arXiv:0806.0848)

- start from Cambridge-Aachen R=0.8 jet
- adjacency: $\Delta R(A,B) > 0.4 - 0.004 \times p_T^{\text{input}}$, do not decluster if fails
- softness: $p_T^{\text{subcluster}} > 0.05 \times p_T^{\text{jet}}$



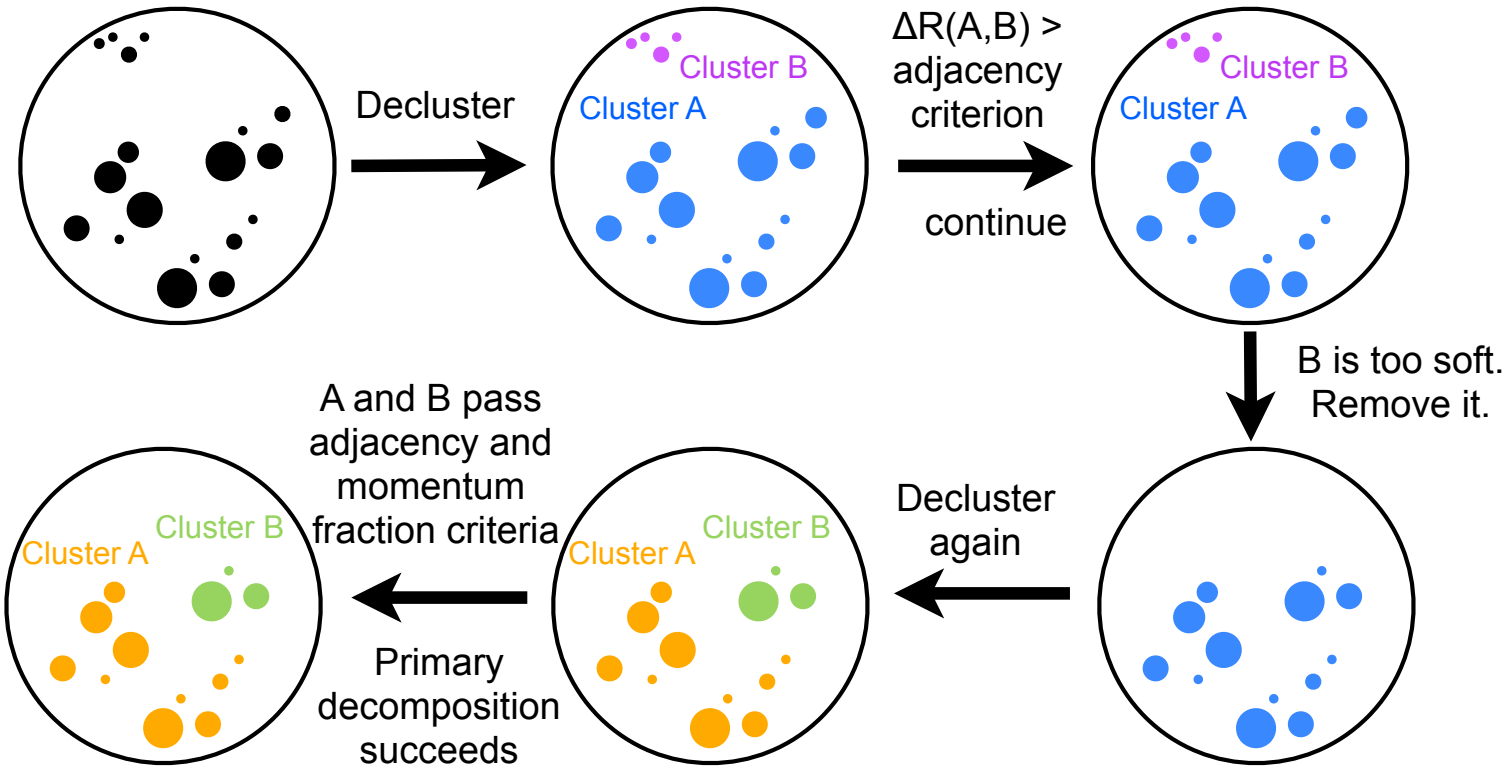
CMS Top Tagger (Kaplan et al., arXiv:0806.0848)

- start from Cambridge-Aachen R=0.8 jet
- adjacency: $\Delta R(A,B) > 0.4 - 0.004 \times p_T^{\text{input}}$, do not decluster if fails
- softness: $p_T^{\text{subcluster}} > 0.05 \times p_T^{\text{jet}}$



CMS Top Tagger (Kaplan et al., arXiv:0806.0848)

- start from Cambridge-Aachen R=0.8 jet
- adjacency: $\Delta R(A,B) > 0.4 - 0.004 \times p_T^{\text{input}}$, do not decluster if fails
- softness: $p_T^{\text{subcluster}} > 0.05 \times p_T^{\text{jet}}$



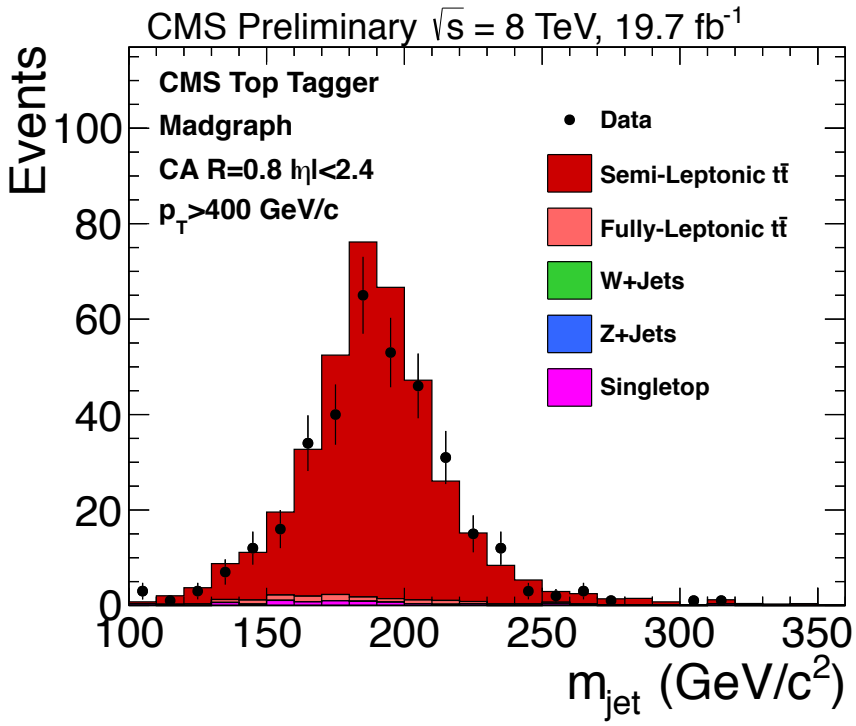
Repeat the procedure separately on cluster A and B

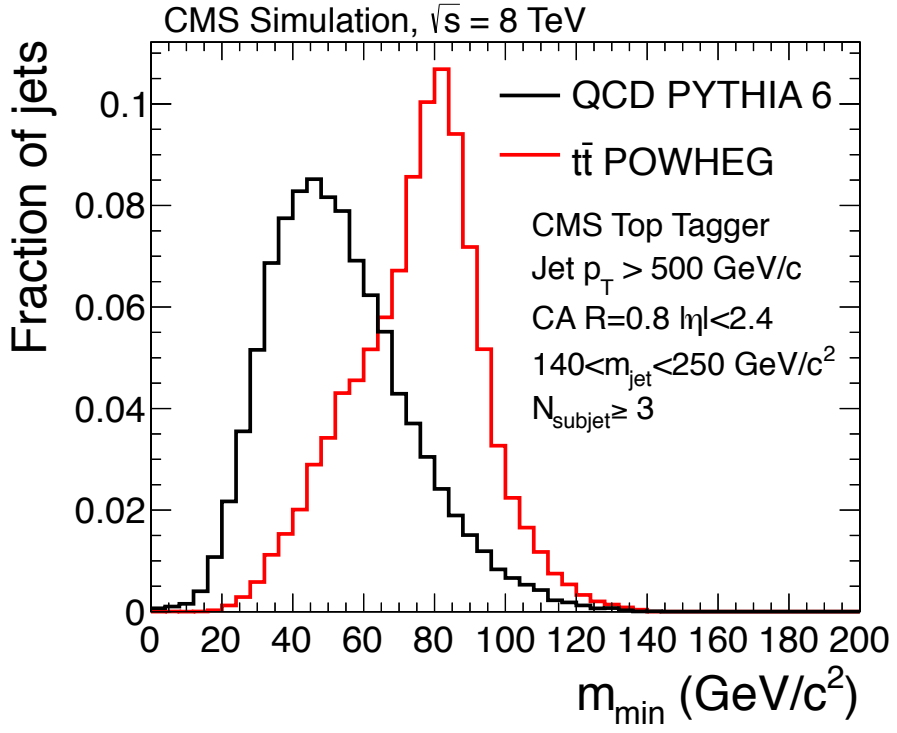
CMS Top Tagger in data

Additional selection:

- $N_{\text{subjet}} \geq 3$
- $m_{\text{min}} = \min(m_{12}, m_{13}, m_{23}) > 50$ GeV
- $140 < m_{\text{jet}} < 250$ GeV

Fully commissioned in CMS in a boosted semileptonic $t\bar{t}$ selection.

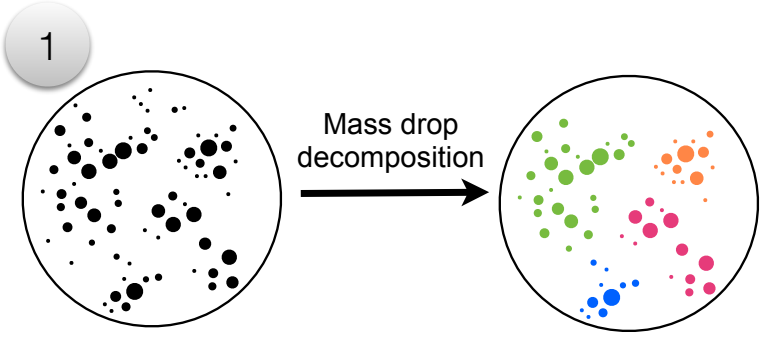




CMS-PAS-JME-13-007

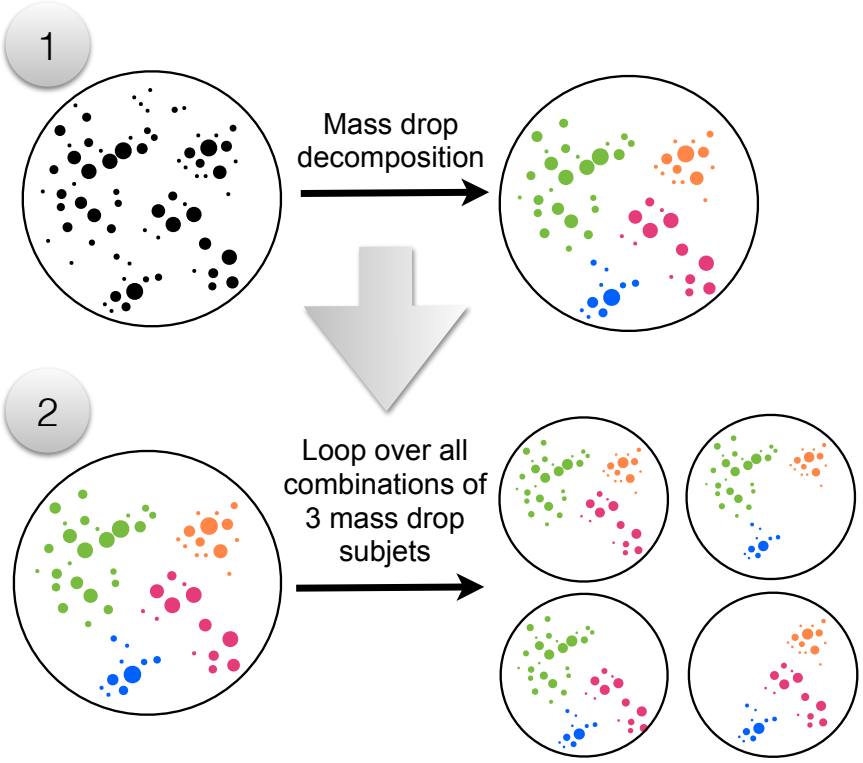
HEP Top Tagger (Plehn et al., arXiv:1006.2833)

- start from Cambridge-Aachen R=1.5 jet



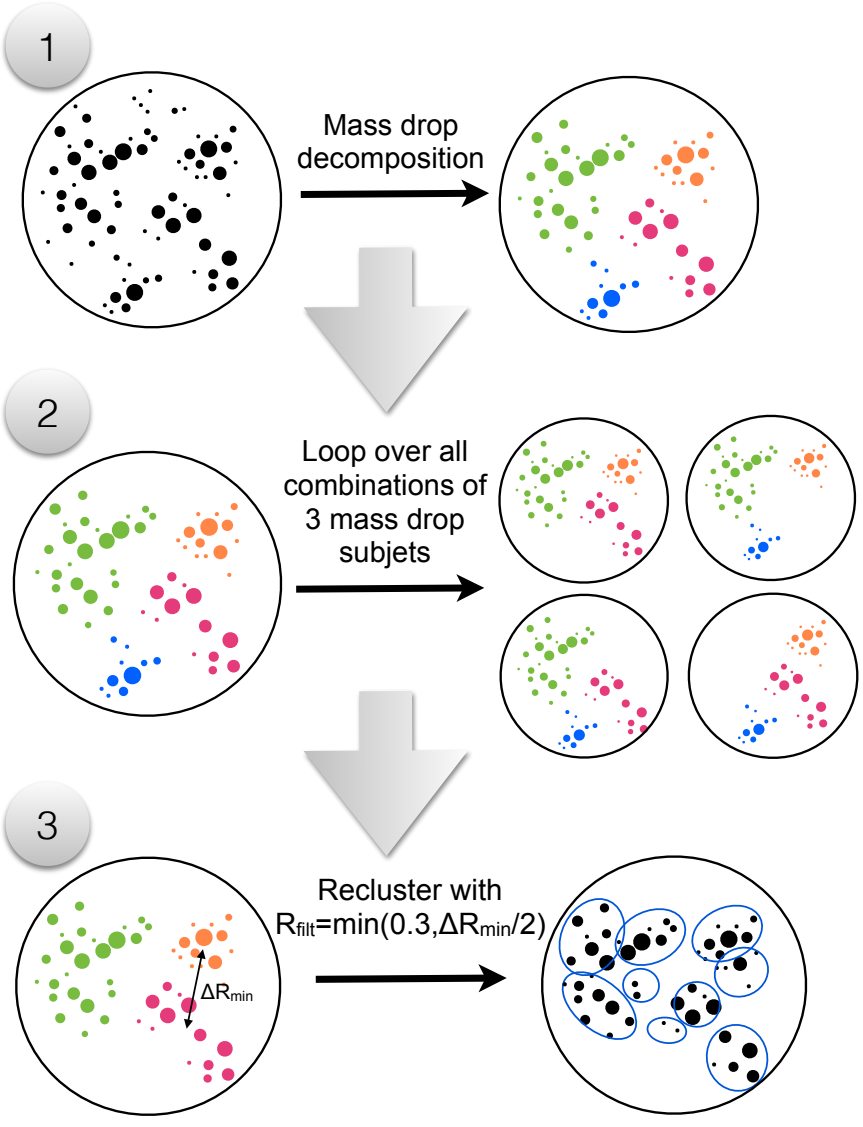
HEP Top Tagger (Plehn et al., arXiv:1006.2833)

- start from Cambridge-Aachen R=1.5 jet



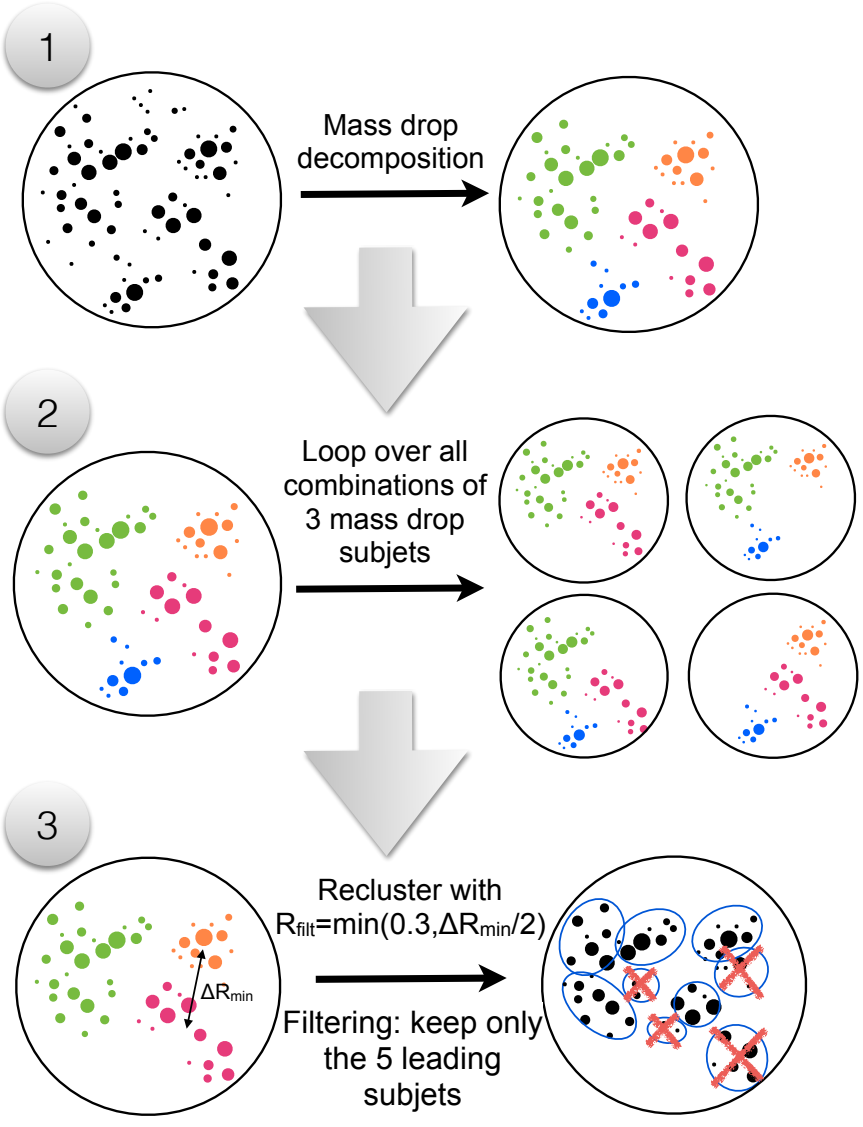
HEP Top Tagger (Plehn et al., arXiv:1006.2833)

- start from Cambridge-Aachen R=1.5 jet



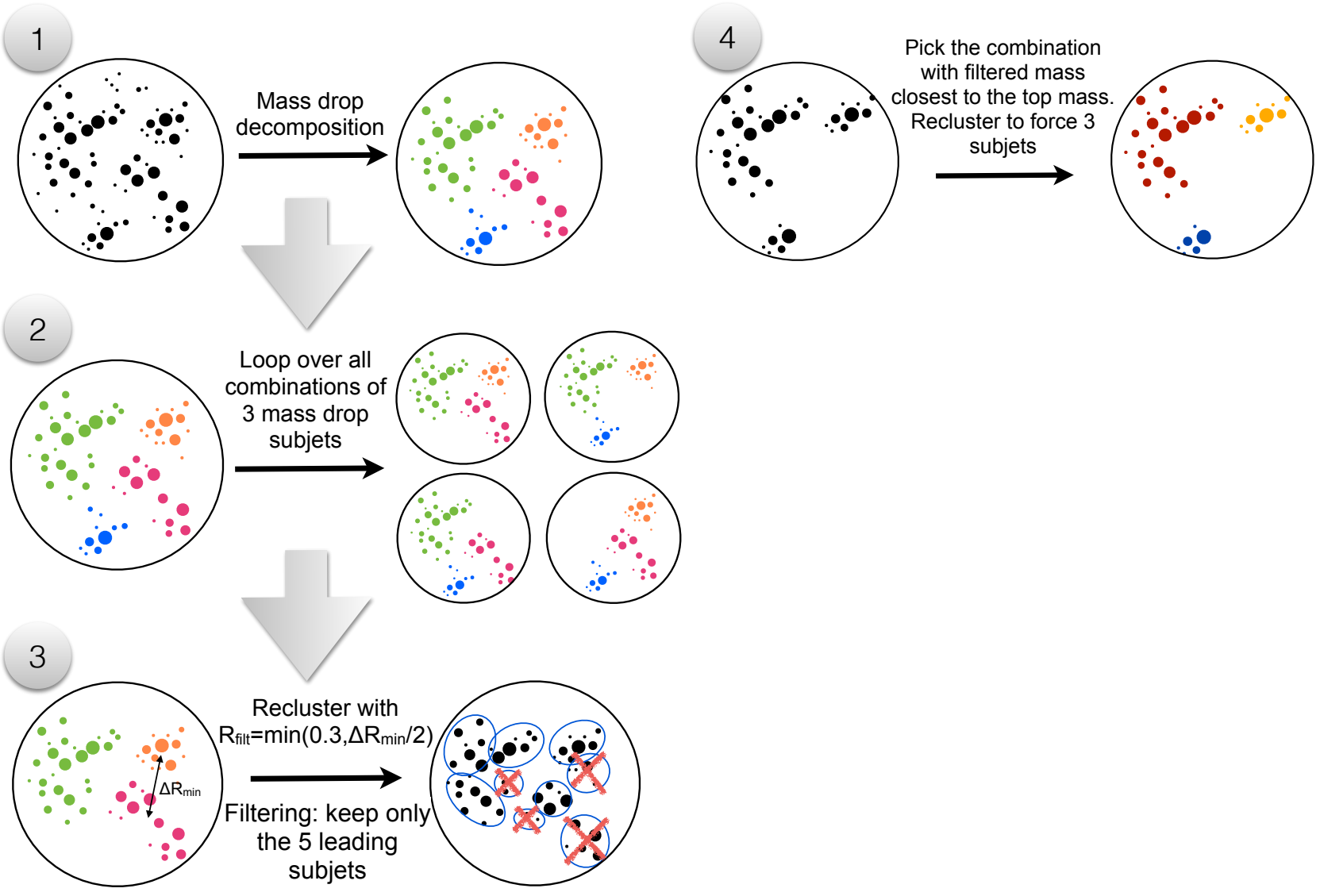
HEP Top Tagger (Plehn et al., arXiv:1006.2833)

- start from Cambridge-Aachen R=1.5 jet



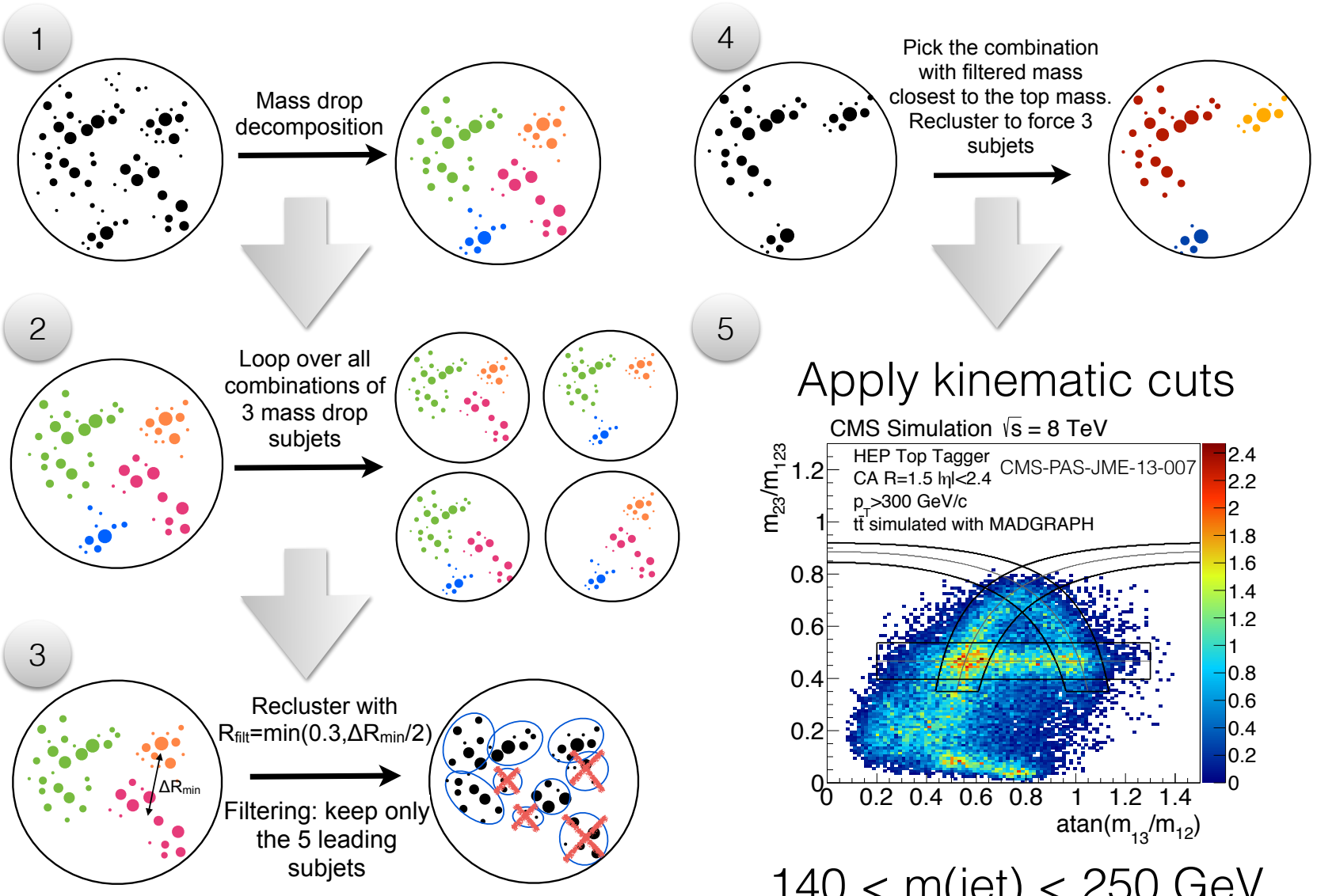
HEP Top Tagger (Plehn et al., arXiv:1006.2833)

- start from Cambridge-Aachen R=1.5 jet



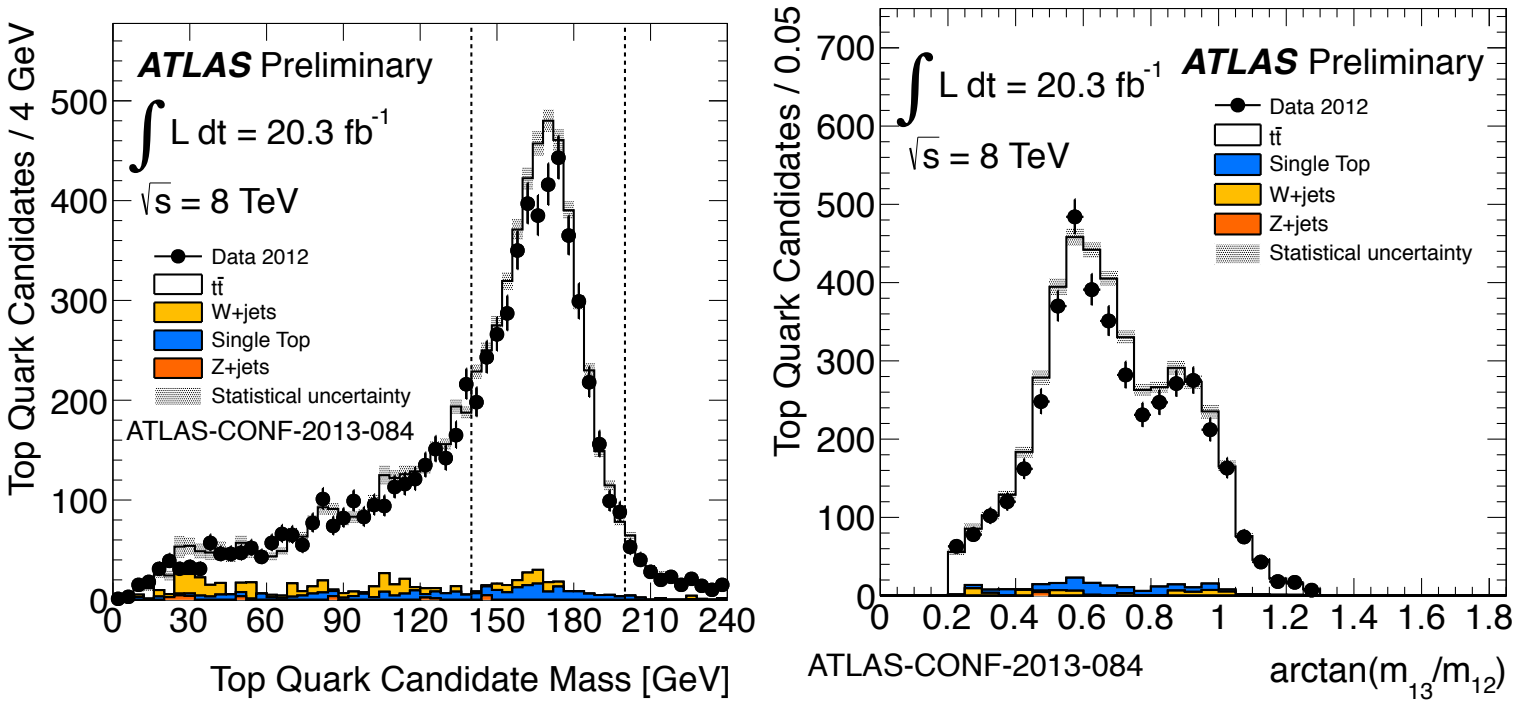
HEP Top Tagger (Plehn et al., arXiv:1006.2833)

- start from Cambridge-Aachen R=1.5 jet



HEP Top Tagger in data

Fully commissioned in both ATLAS and CMS in a boosted semileptonic $t\bar{t}$ selection.



Tagger's efficiency drops at very high $p_T \rightarrow$ cone is too large $\rightarrow$ new improved Multi-R HepTopTagger:

- run tagger on different cone sizes R (0.5-1.5)
- find optimal cone size:$R_{\text{min}} = \min(R' | \text{abs}(m_{\text{jet}}(R = R') - m_{\text{jet}}(R = 1.5)) < 0.2 m_{\text{jet}}(R = 1.5))$
- additionally use $R_{\text{min}}(p_T) - R_{\text{min,expected}}(p_T)$ as tagging variable

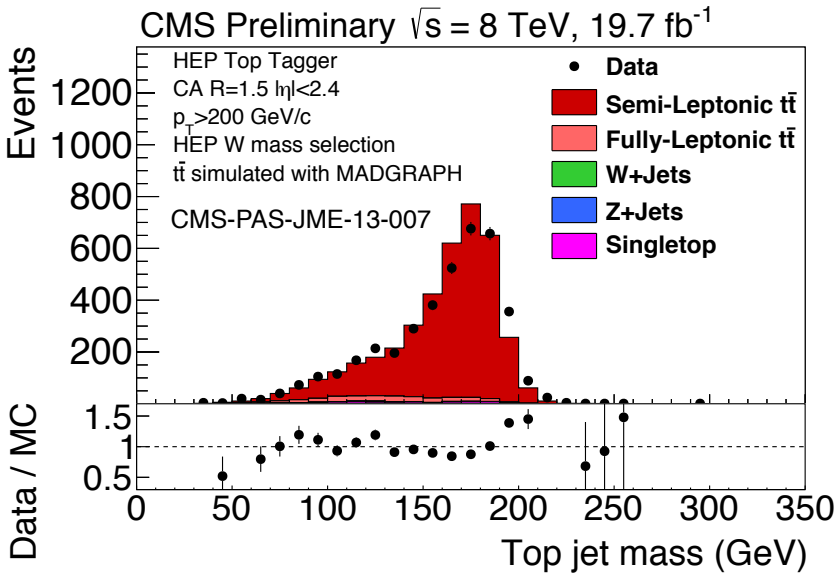
Calibration and data driven corrections

CMS: derive Data/MC eff. correction factor
Tag&probe on $t\bar{t}$ semileptonic sample:

- ''tag'' on leptonic side (iso μ, b-jet)
- ''probe'' effic. on the hadr. side

Scale factors derived as a function of:

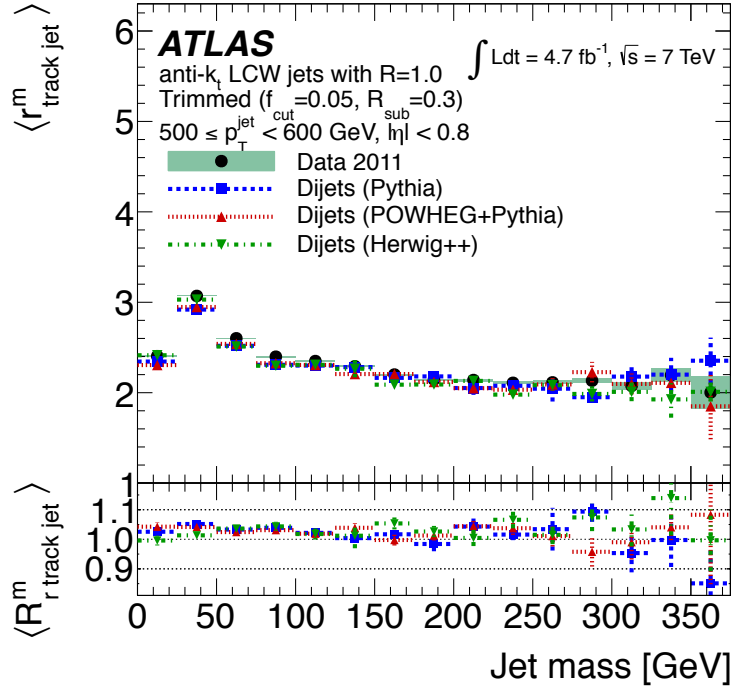
- p_T, η, generators



ATLAS: (arXiv:1306.4945)
HEP Tagger: calibrate subjets for different R
Large cone jets: validate jet mass scale using:

- $r^m_{\text{track jet}} = \frac{m^{\text{jet}}}{m^{\text{track jet}}}$
- $R^m_{r\ \text{track jet}} = \frac{r^{m,\text{data}}_{\text{track jet}}}{r^{m,\text{MC}}_{\text{track jet}}}$

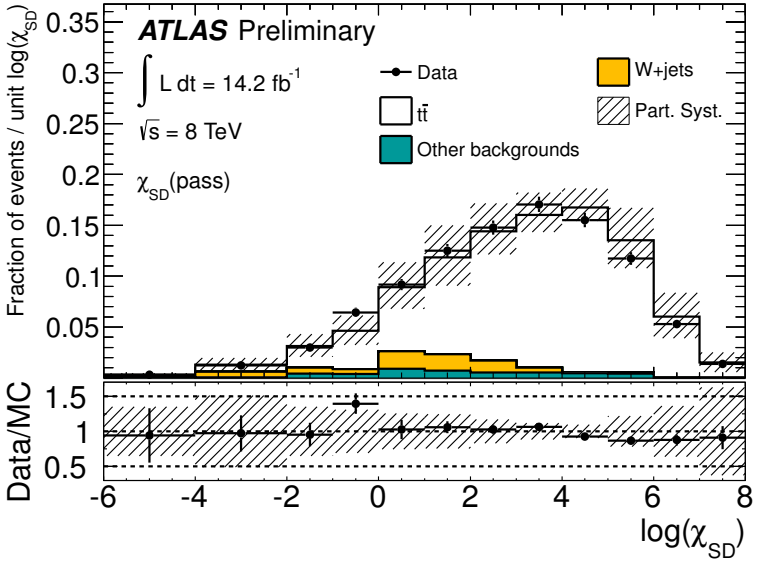
$\rightarrow$ derive mass scale and JES uncertainties

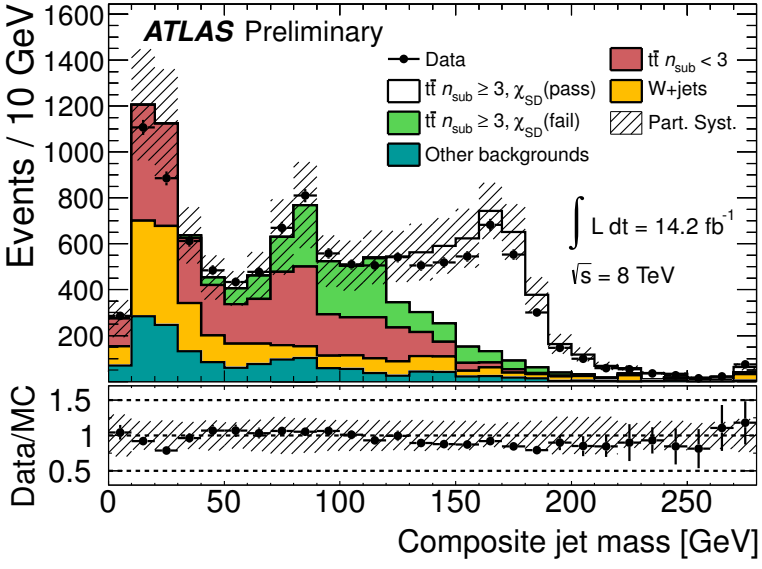


Shower deconstruction

- decluster small *microjets* (R=0.1-0.3) from a large cone jet
- assume each microjet (momentum p) comes from a particle (top decay products, ISR parton shower, etc.)
- derive probability a microjet configuration $\{p\}_N$ comes from a particular decay chain (shower history)
- probability obtained by combining splitting probabilities from shower history
- define discriminator χ as S/B probability quotient

$$\chi(\{p\}_N) = \frac{P(\{p\}_N|\text{signal})}{P(\{p\}_N|\text{background})}$$





ATLAS-CONF-2014-003

Further jet substructure techniques

b-tagging in boosted topologies

_b_-tagging: essential tool against QCD

Algorithms use information from:

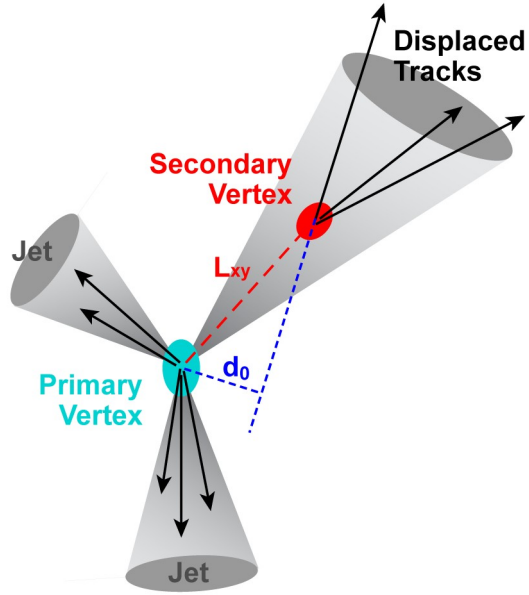
- tracks, esp. impact parameter (IP)
- secondary vertices (SV)

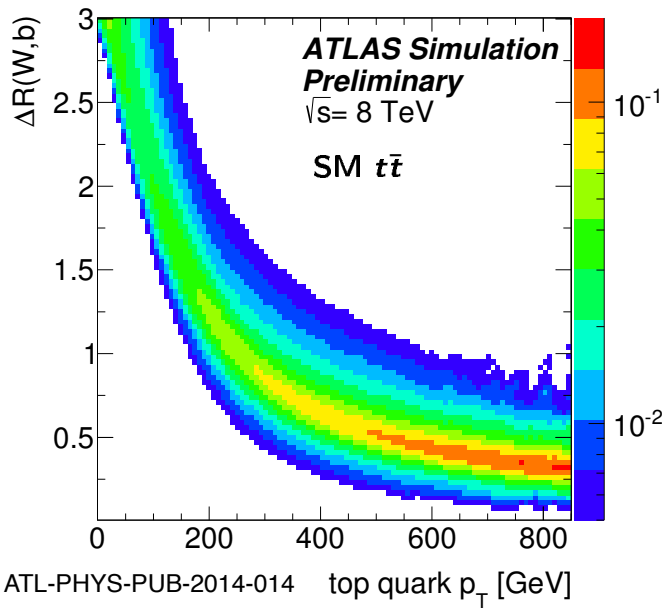
Combine information in a discriminator

- ATLAS MV1: neural network based
- CMS Combined Secondary Vertex (CSV): likelihood based

In boosted regime:

- small separation of decay products
- light flavor contamination
- *busy* environment
- degraded performance

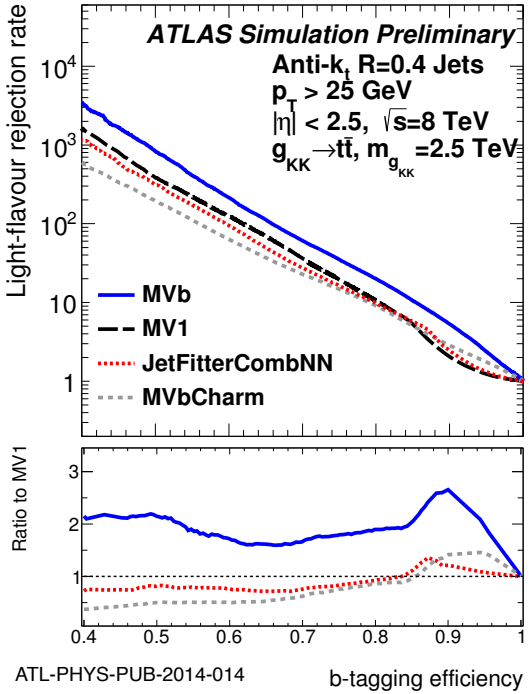


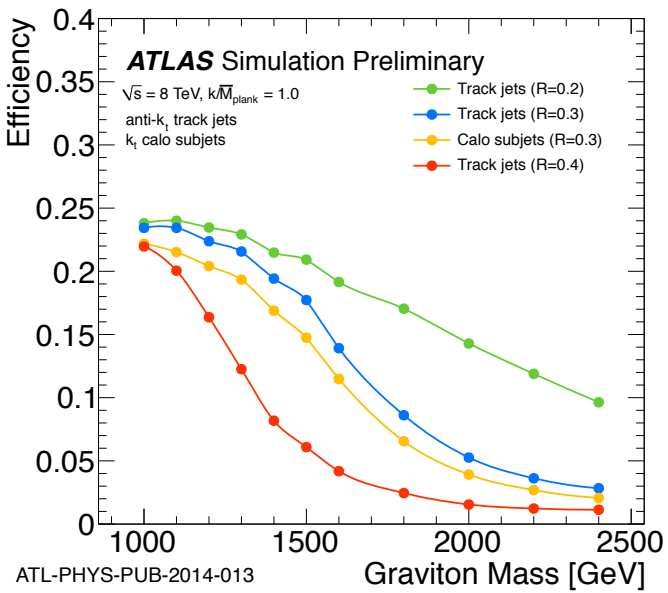


b-tagging in boosted topologies – ATLAS

An improved tagger:

- add new variables with more discrimination power in boosted regimes
  (mean ΔR(trk,jet), 3rd highest d_0 significance)
- retrain NN with boosted $t\bar{t}$ decays
- **new MVb** 2× better than **MV1** in the boosted regime





Use different input jets

- small-R jets: resolve better boosted decay products
- use track-jets instead of calo-jets: better jet direction resolution

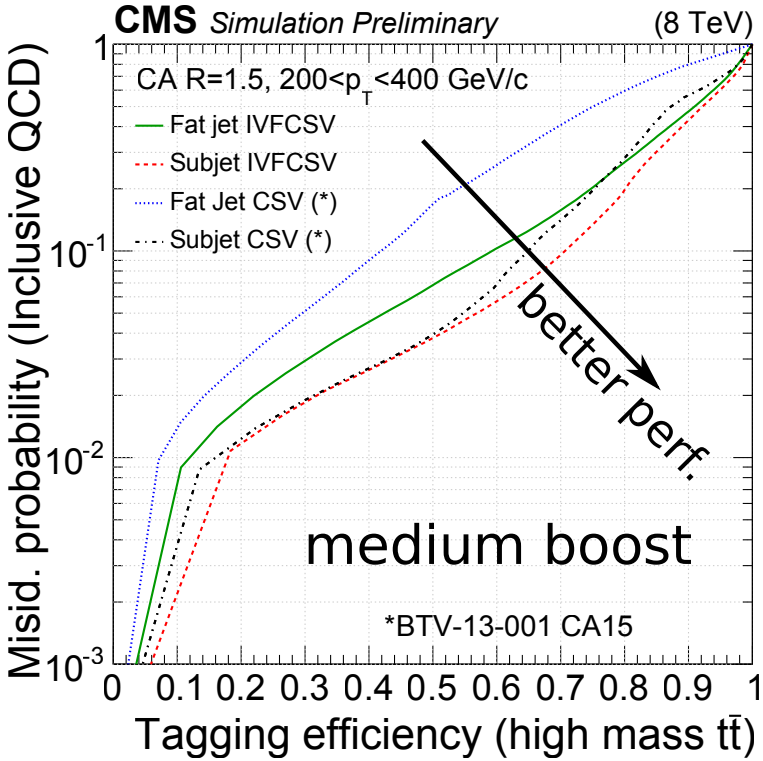
b-tagging in boosted topologies – CMS

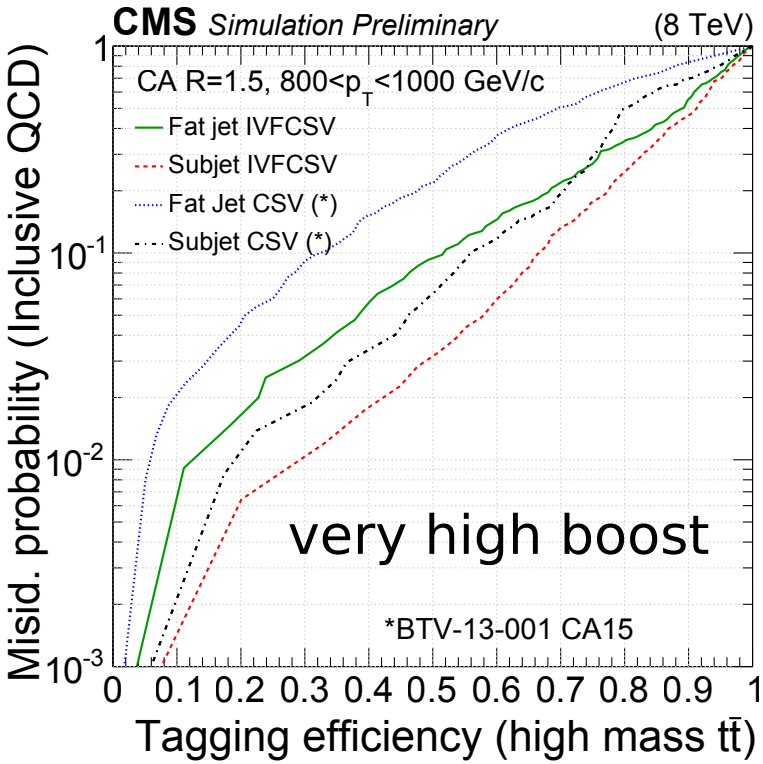
Two different approaches

- fat jet *b*-tag: apply CSV directly to fat jets (**green** and **blue**)
- subjet *b*-tag: apply CSV to subjets of fat jets (**red** and **black**)

Several algorithm improvements (improved taggers in **green** and **red**)

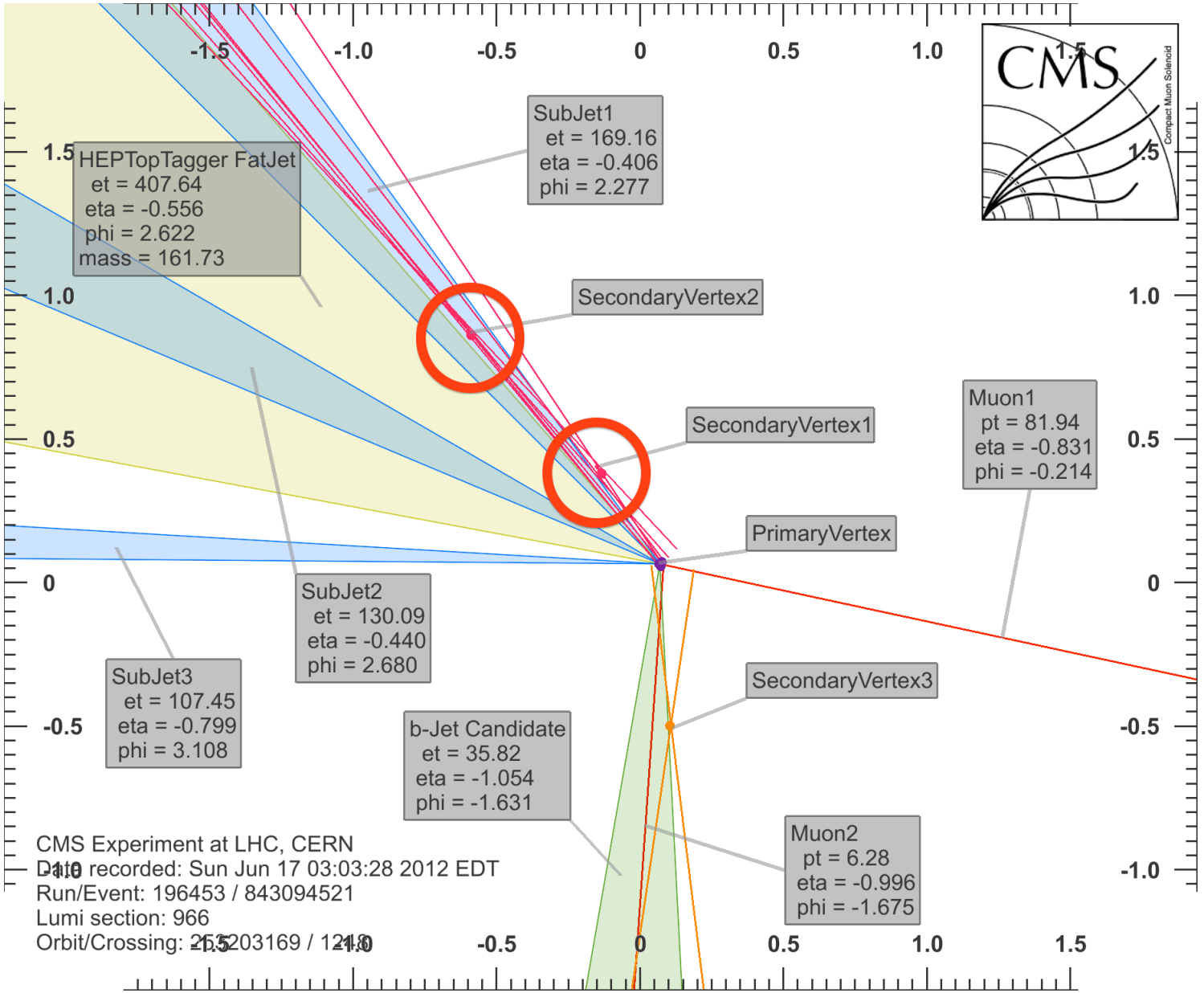
- Inclusive Vertex Finder (IVF)
- use tracks linked to charged constituents of particle-flow jets (instead of fixed R cone) + other CSV improvements





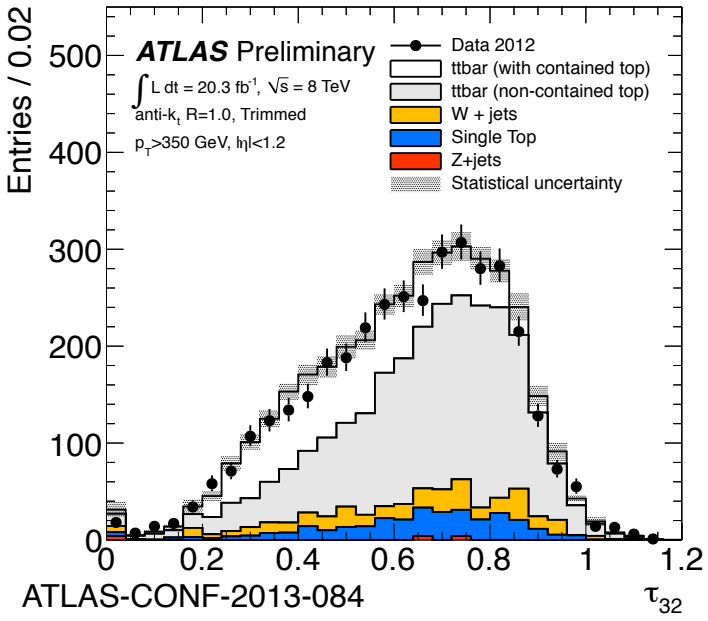
CMS DP-2014/031

subjet *b*-tagging in action in CMS (data)



CMS-PAS-BTV-13-001

N-subjettiness and k_t splitting scale

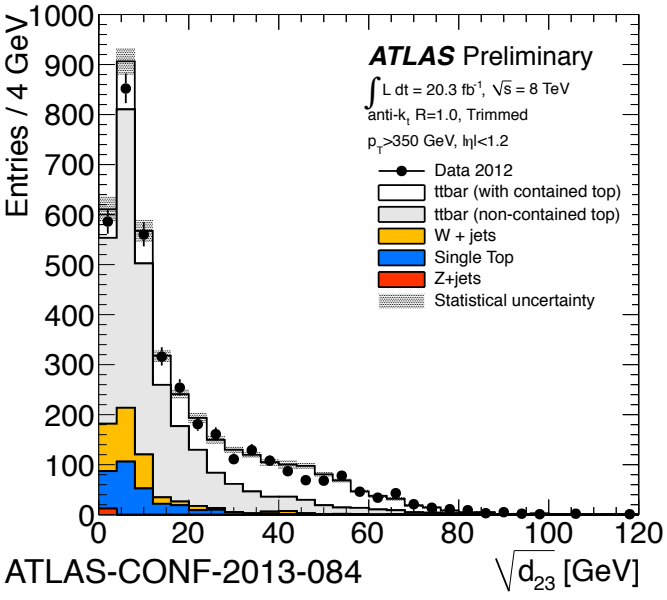


N-subjettiness τ_N (ATLAS+CMS):

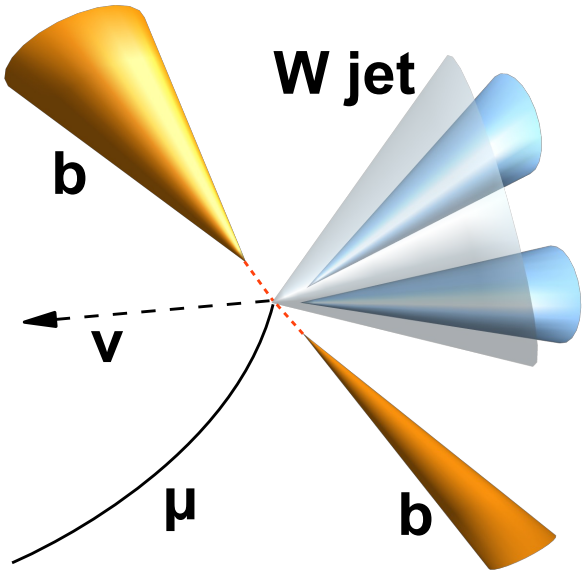
- how well jets can be described as containing N or fewer k_t subjets
- $\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$
  k: constituent index, $d_0 = \sum_k p_{T,k} R_0$
- Roughly: p_T weighted average of minimum ΔR(constituent, subjet axis)$/R_{jet}$
- for tops τ_3/τ_2 and τ_2/τ_1 are relevant

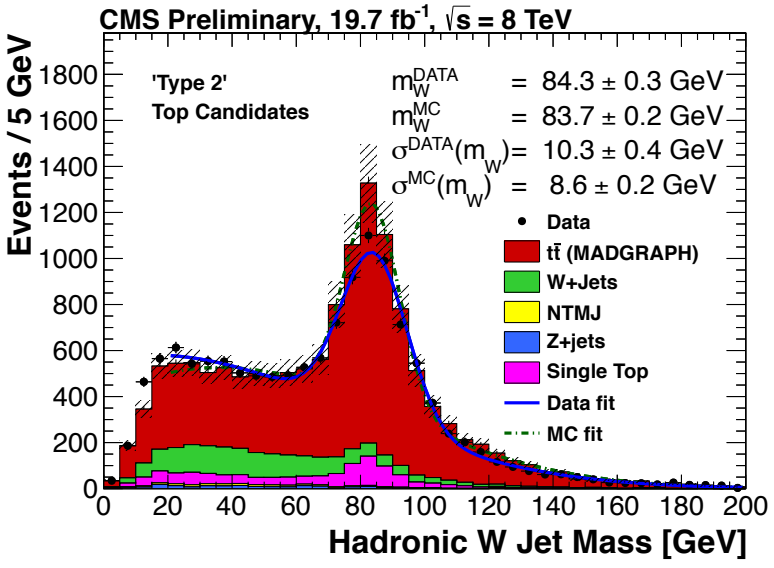
k_t splitting scales (ATLAS):

- recluster jet constituents with k_t algorithm
- use last combined jets to define the splitting scale: $\sqrt{d_{i,j}} = \min(p_{Ti}, p_{Tj}) \times \Delta R_{i,j}$
- $\sqrt{d_{1,2}} \rightarrow$ last step, $\sqrt{d_{2,3}} \rightarrow$ second to last



The semi-resolved case and *W*-tagging





CMS-PAS-JME-13-007

Medium boost regime

- W and b can be clustered separately
- cluster hadronic W in a single jet (CA R=0.8 in CMS)
- use W-tagging

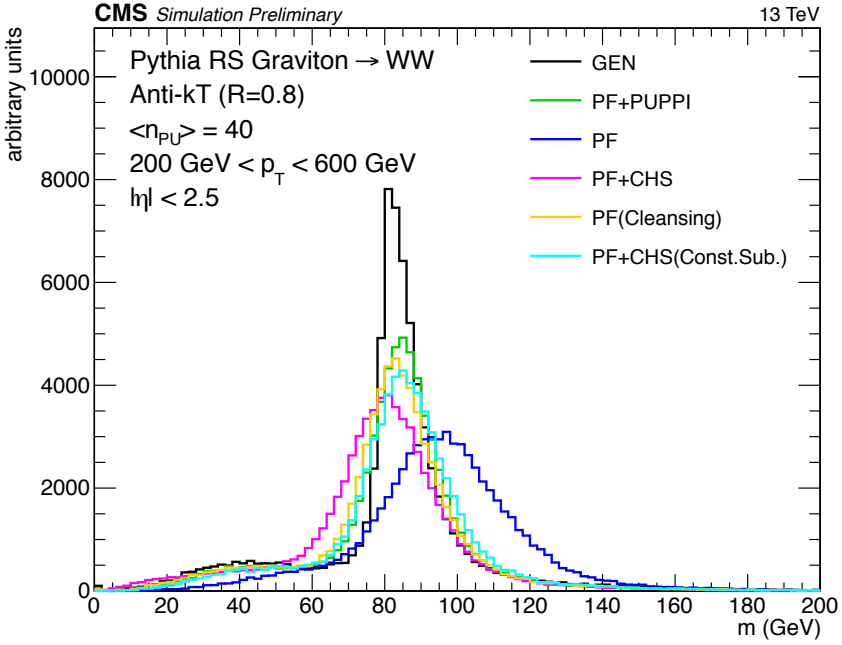
W-tagging

- benchmark topology for jet substructure studies
- treat large W jet with grooming techniques
- apply cut on pruned mass
- some analyses use cut on τ_2/τ_1

Pileup mitigation and other techniques

Pileup per particle identification (CMS)
(PUPPI, JME-14-001,arXiv:1407.6013)

- large-R jets collect lots of pileup
- assign weight to each particle based on:
  - event pileup properties
  - tracking information
- operates on the input of jet algos
- all jet substructure variables affected by it



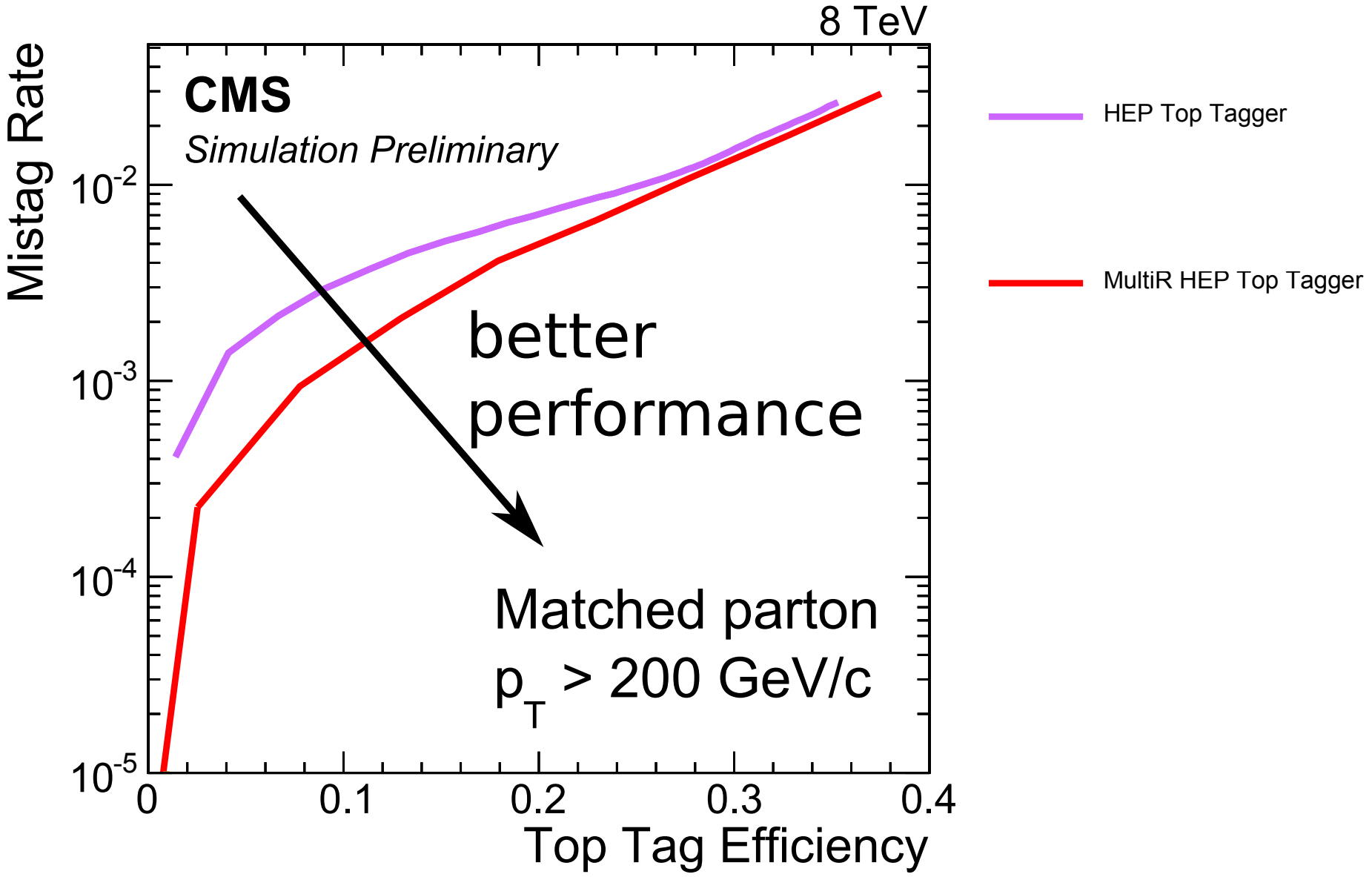
ATLAS: Jet vertex tagger, "Jet Cleansing" (ATLAS-CONF-2014-018)

Jet reclustering (ATLAS)

- Consider the standard R=0.4 jets in an event
- use these jets as an input to recluster anti-k_t jets with larger R=0.8, 1.2
- use mass of the reclustered jets as tagging variable

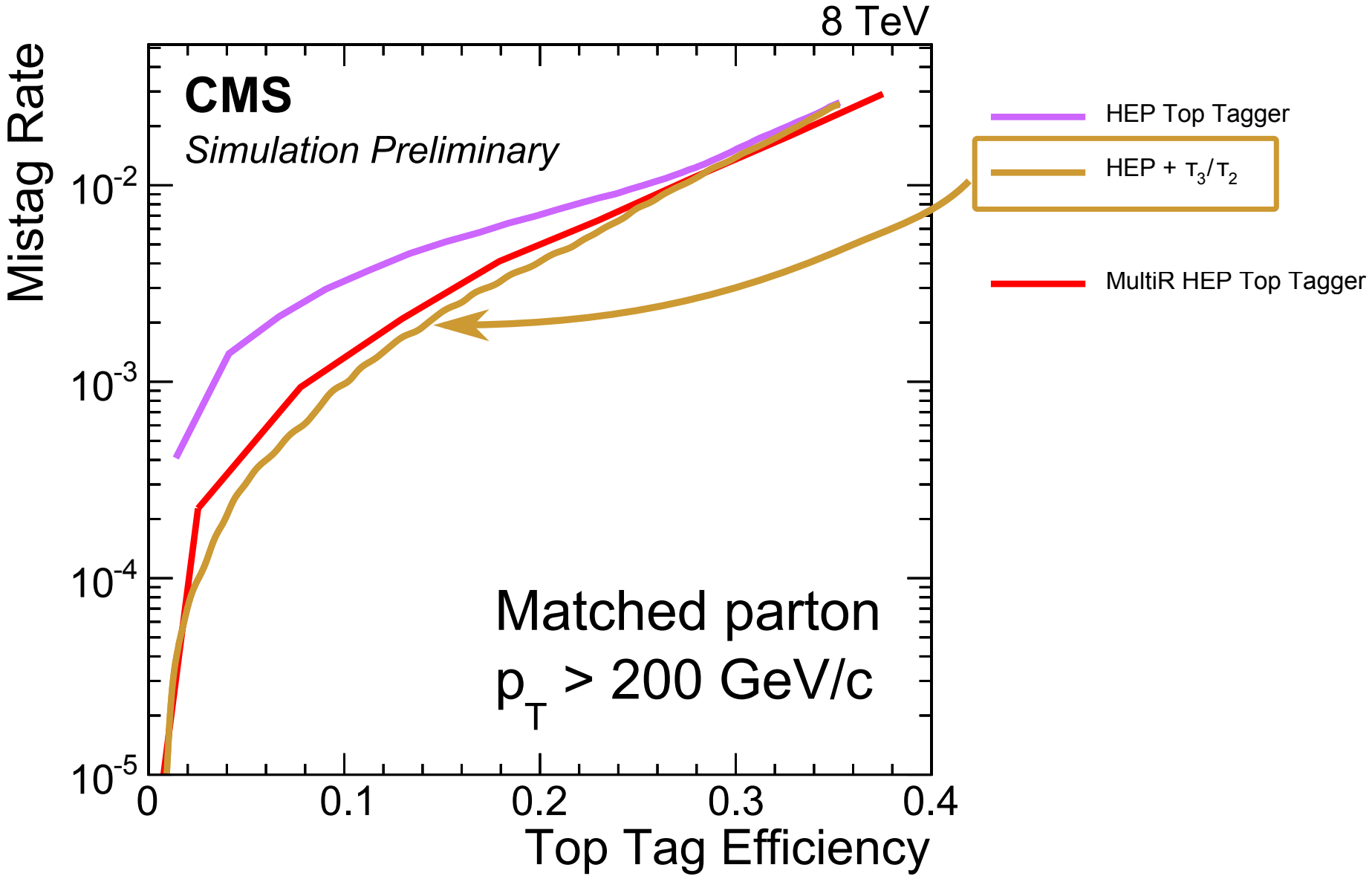
Algorithm comparison

Performance in CMS, $R = 1.5$, $p_{T,\text{match}} > 200$ GeV



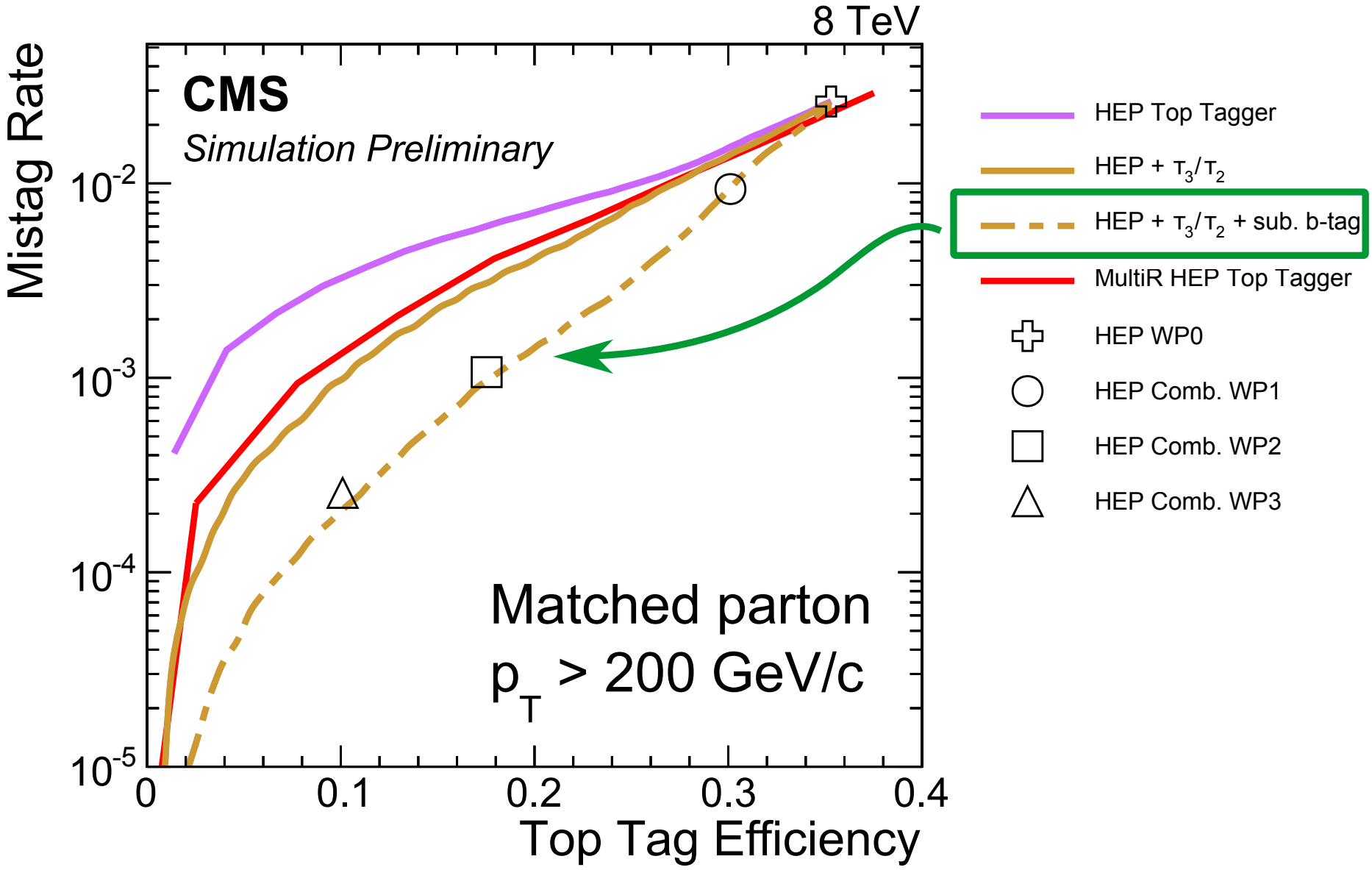
CMS-PAS-JME-13-007

Performance in CMS, $R = 1.5$, $p_{T,\text{match}} > 200$ GeV



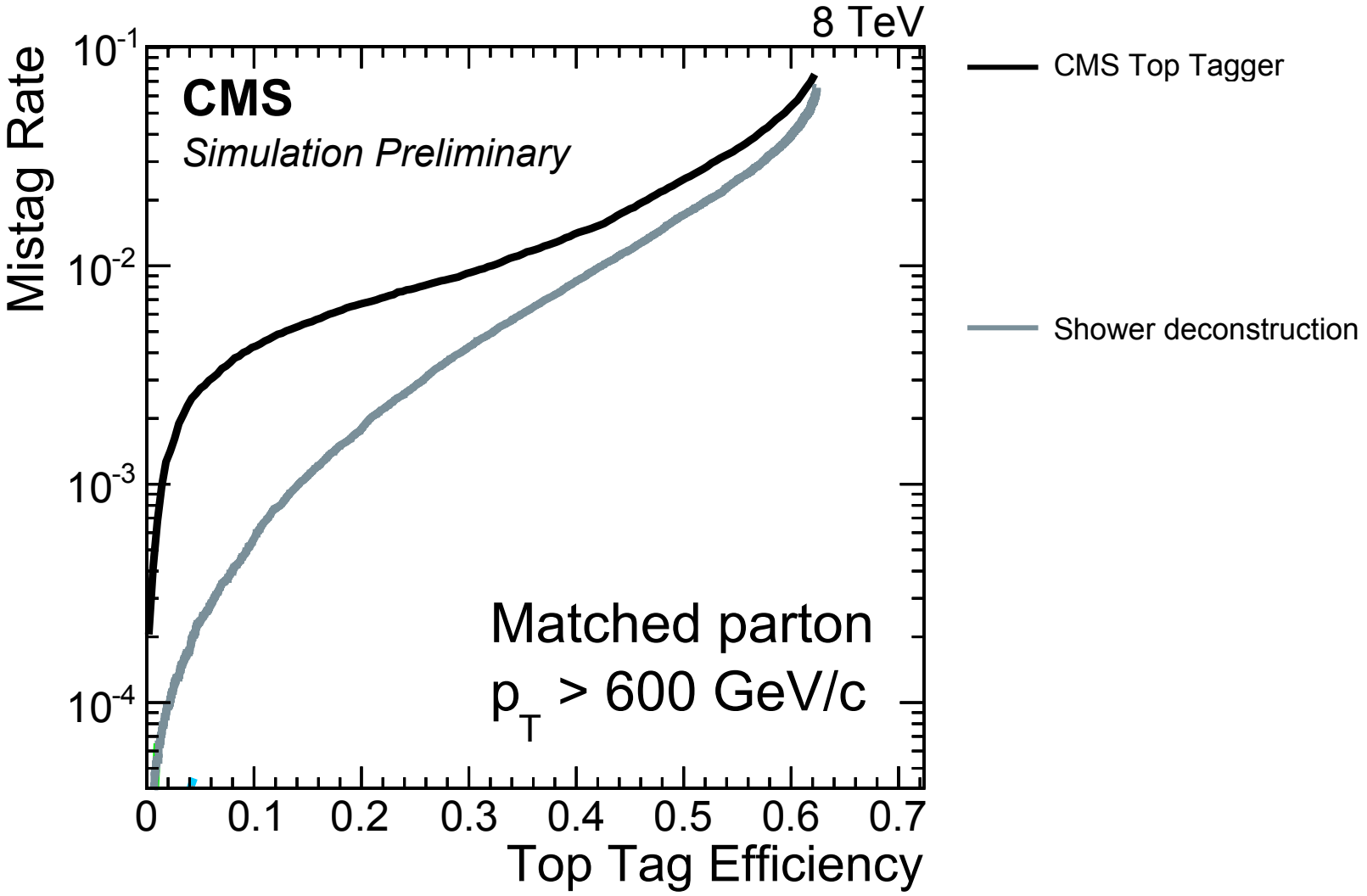
CMS-PAS-JME-13-007

Performance in CMS, $R = 1.5$, $p_{T,\text{match}} > 200$ GeV



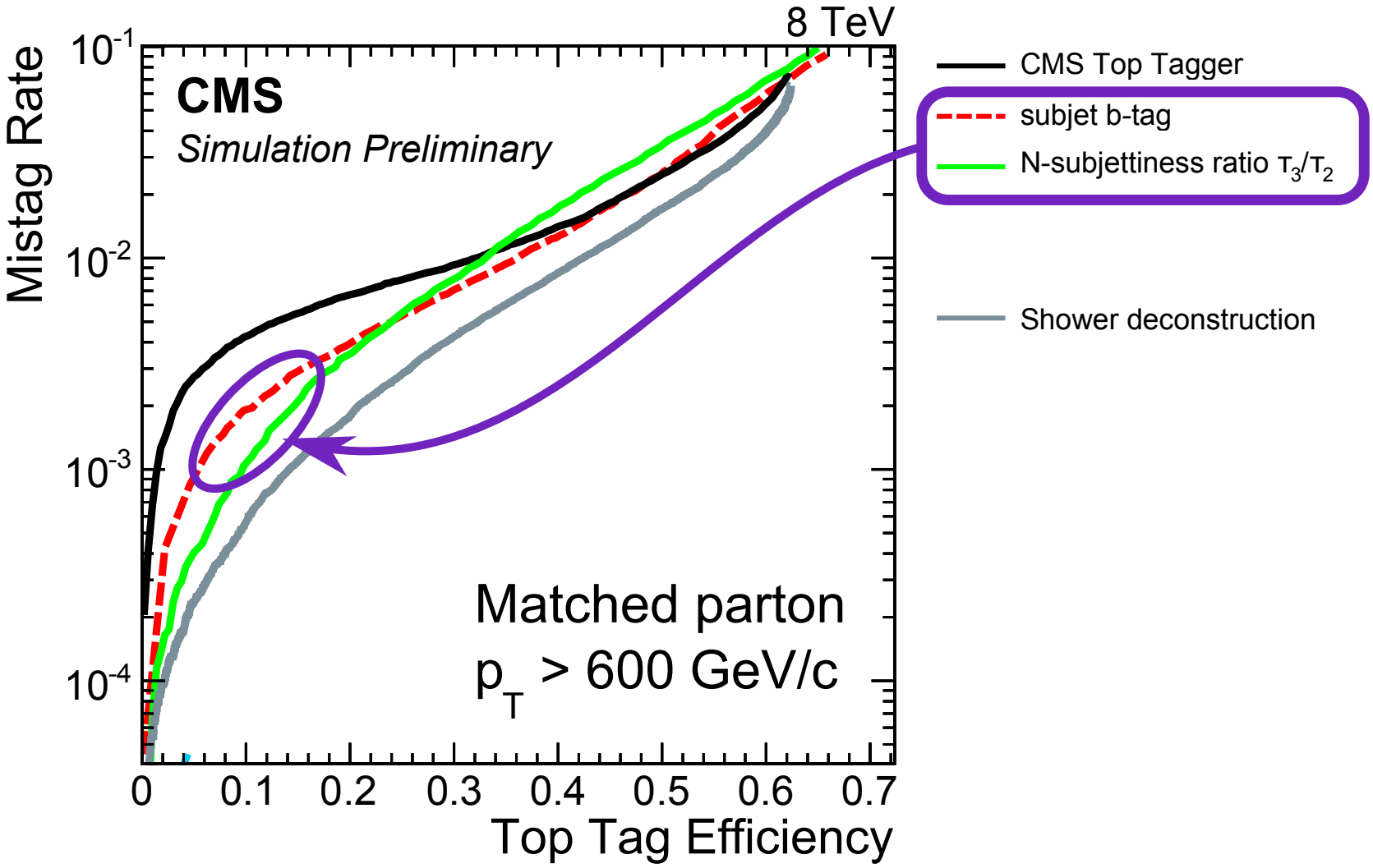
CMS-PAS-JME-13-007

Performance in CMS, $R = 0.8$, $p_{T,\text{match}} > 600$ GeV



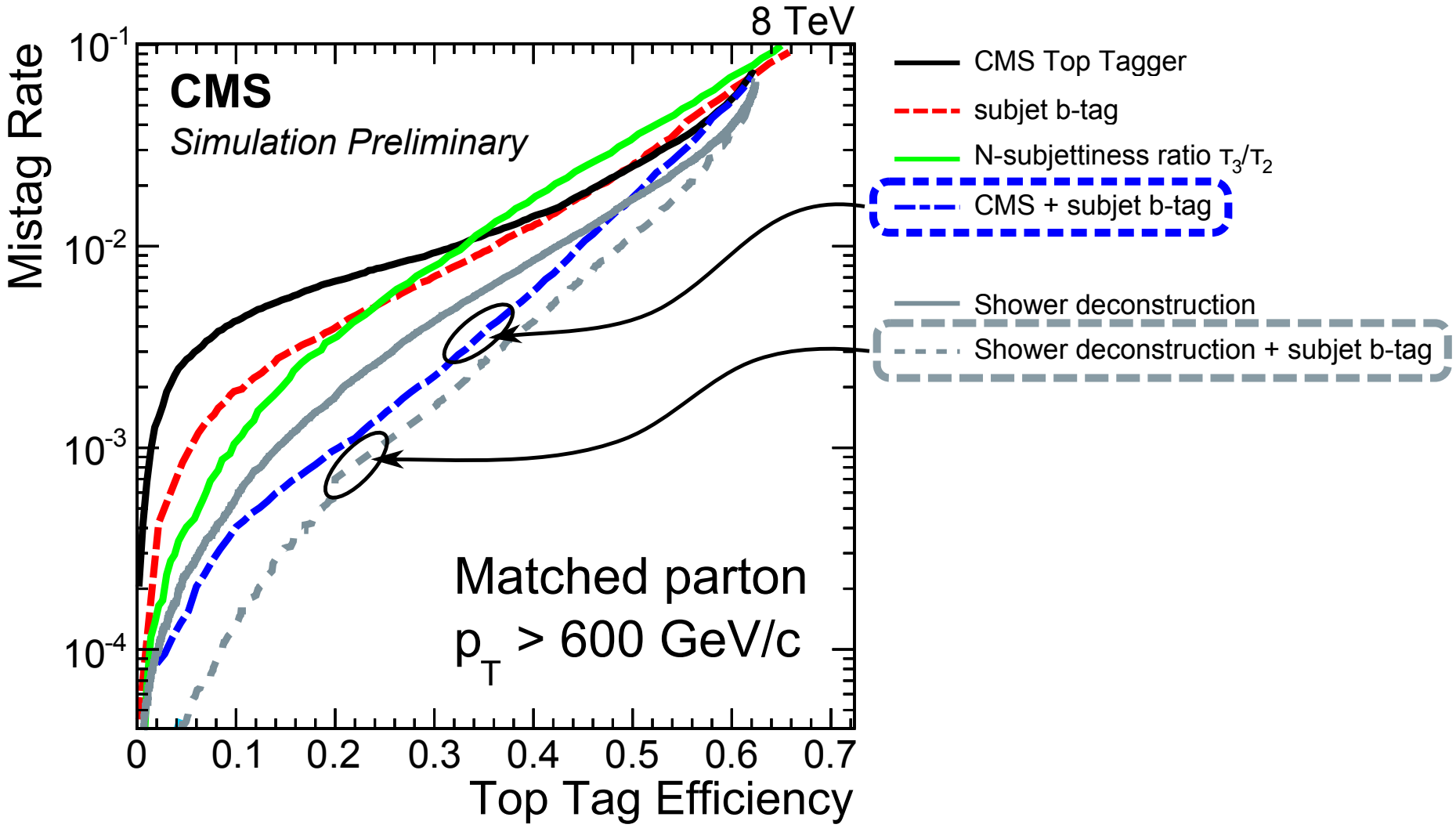
CMS-PAS-JME-13-007

Performance in CMS, $R = 0.8$, $p_{T,\text{match}} > 600$ GeV



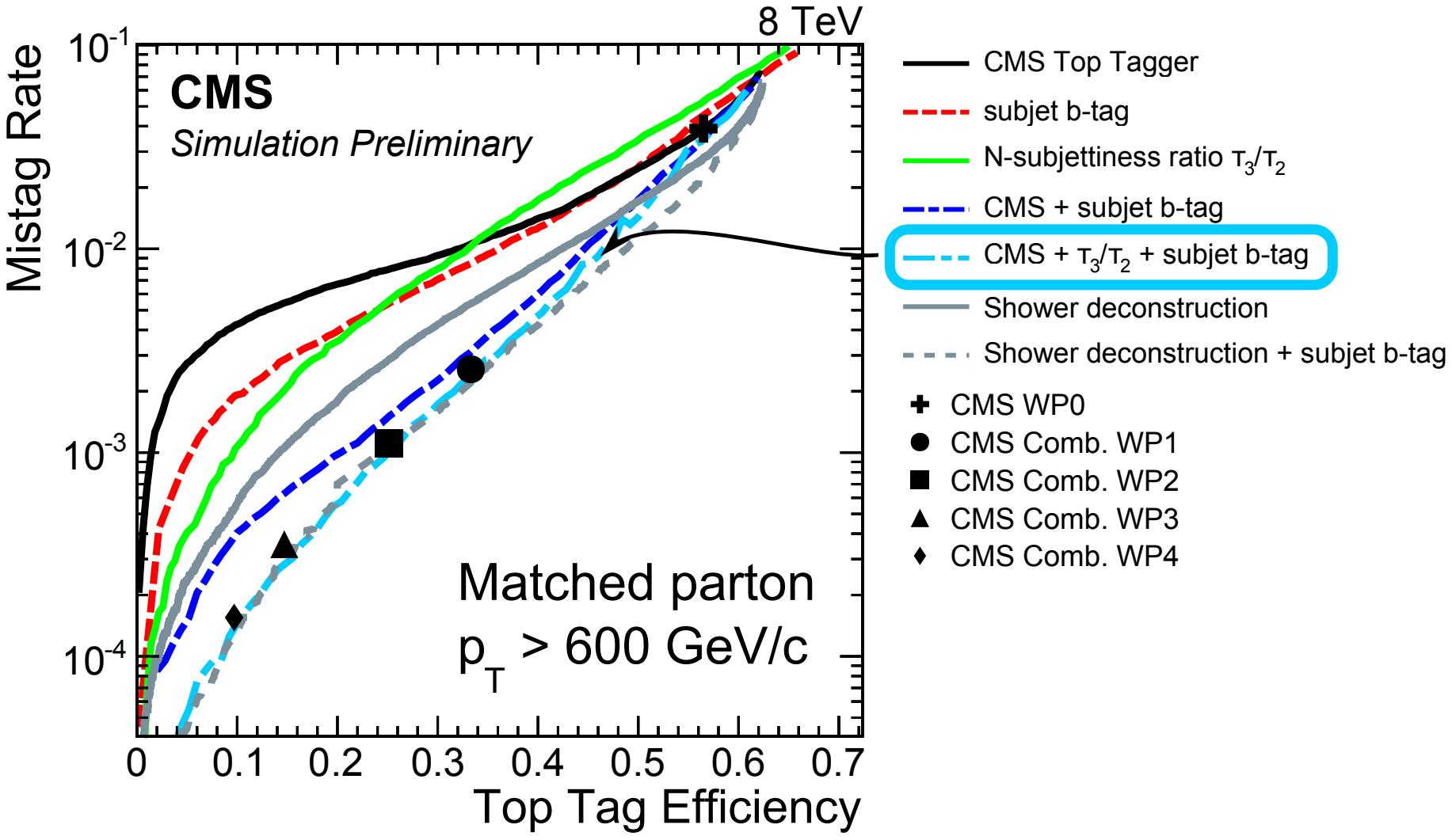
CMS-PAS-JME-13-007

Performance in CMS, $R = 0.8$, $p_{T,\text{match}} > 600$ GeV



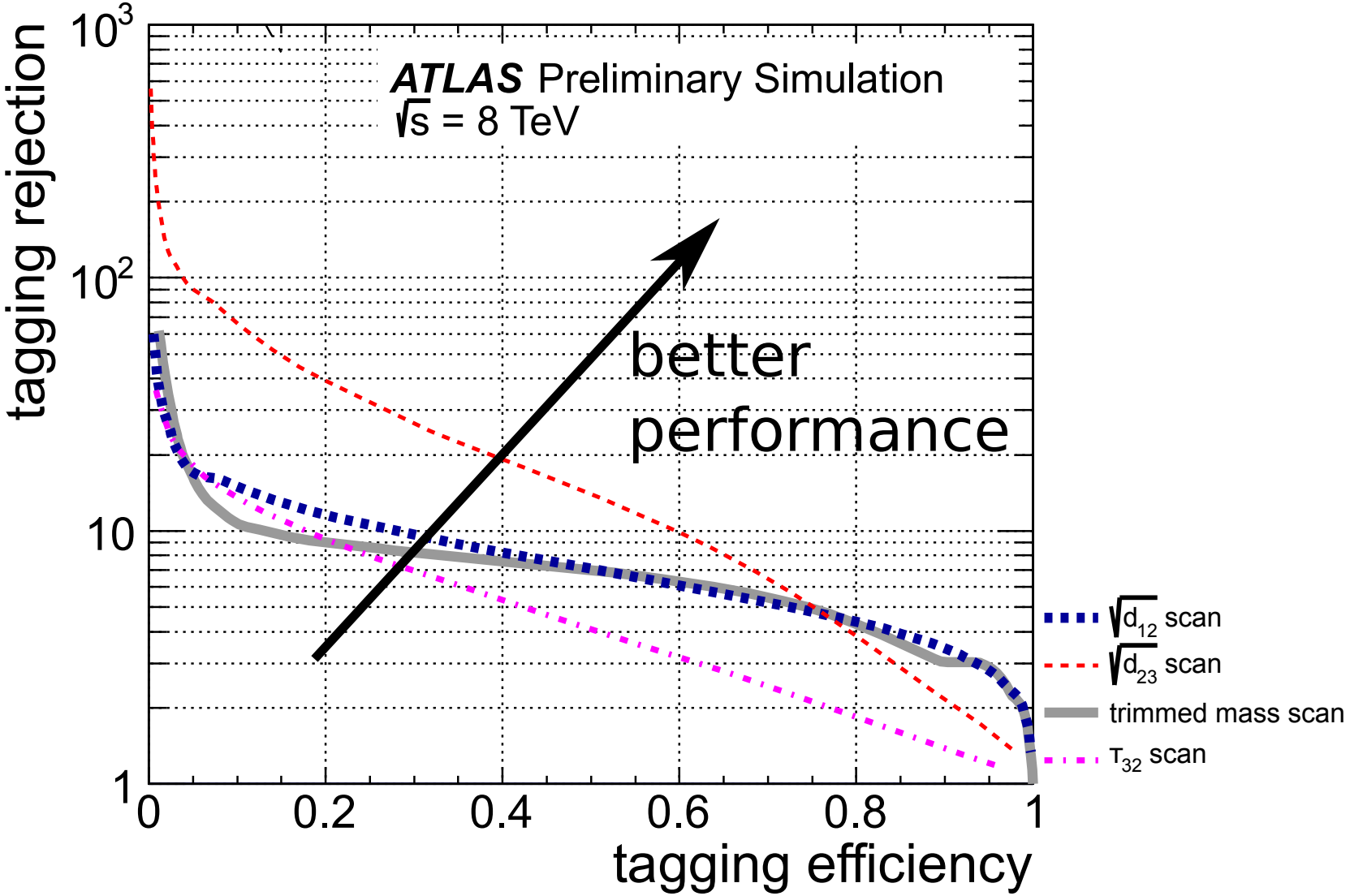
CMS-PAS-JME-13-007

Performance in CMS, $R = 0.8$, $p_{T,\text{match}} > 600$ GeV



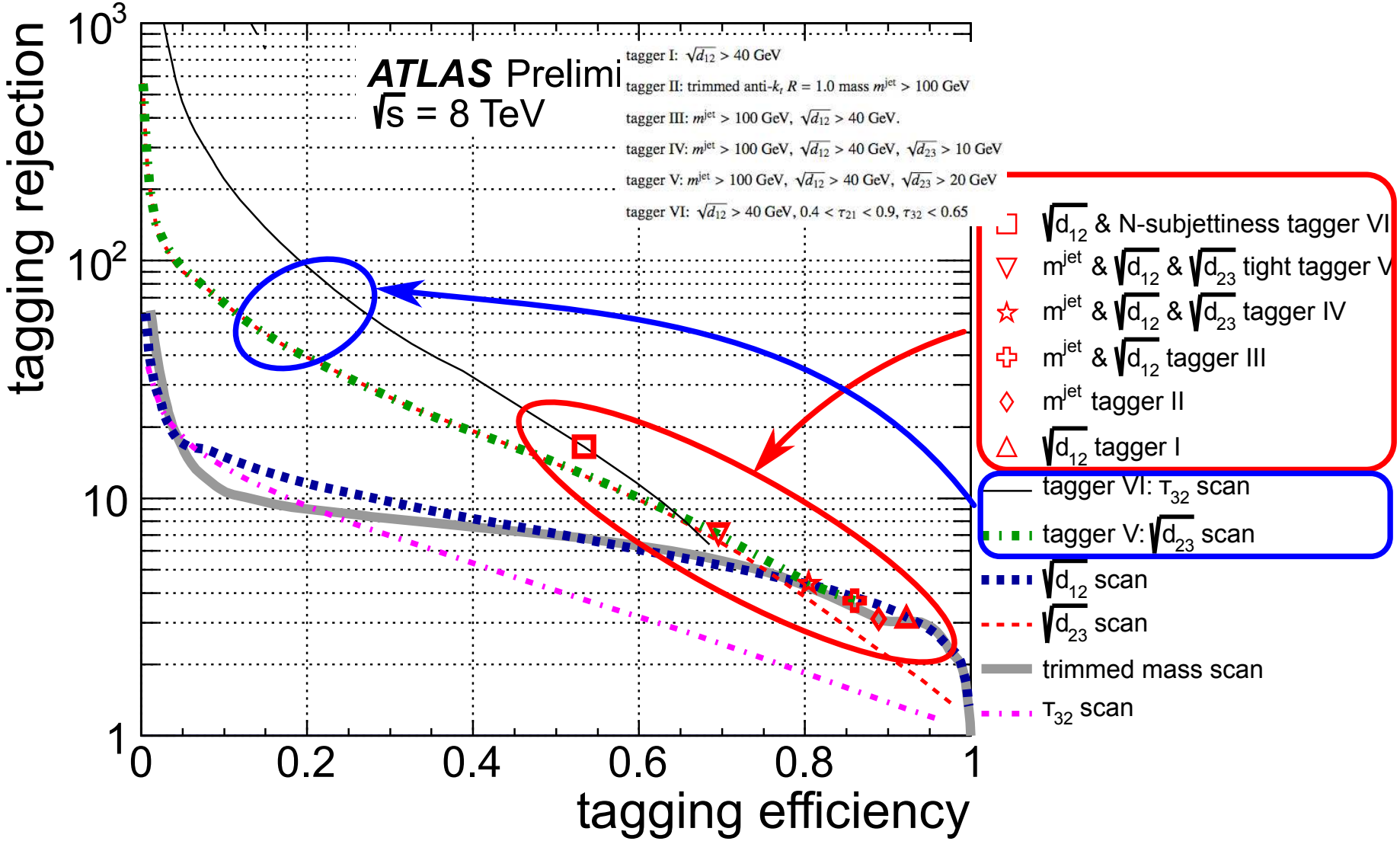
CMS-PAS-JME-13-007

Performance comparison in ATLAS: trimm. a-k_t R=1.0, CA R=1.2 (HEPTT), $p_{T,\text{jet}} > 550$ GeV



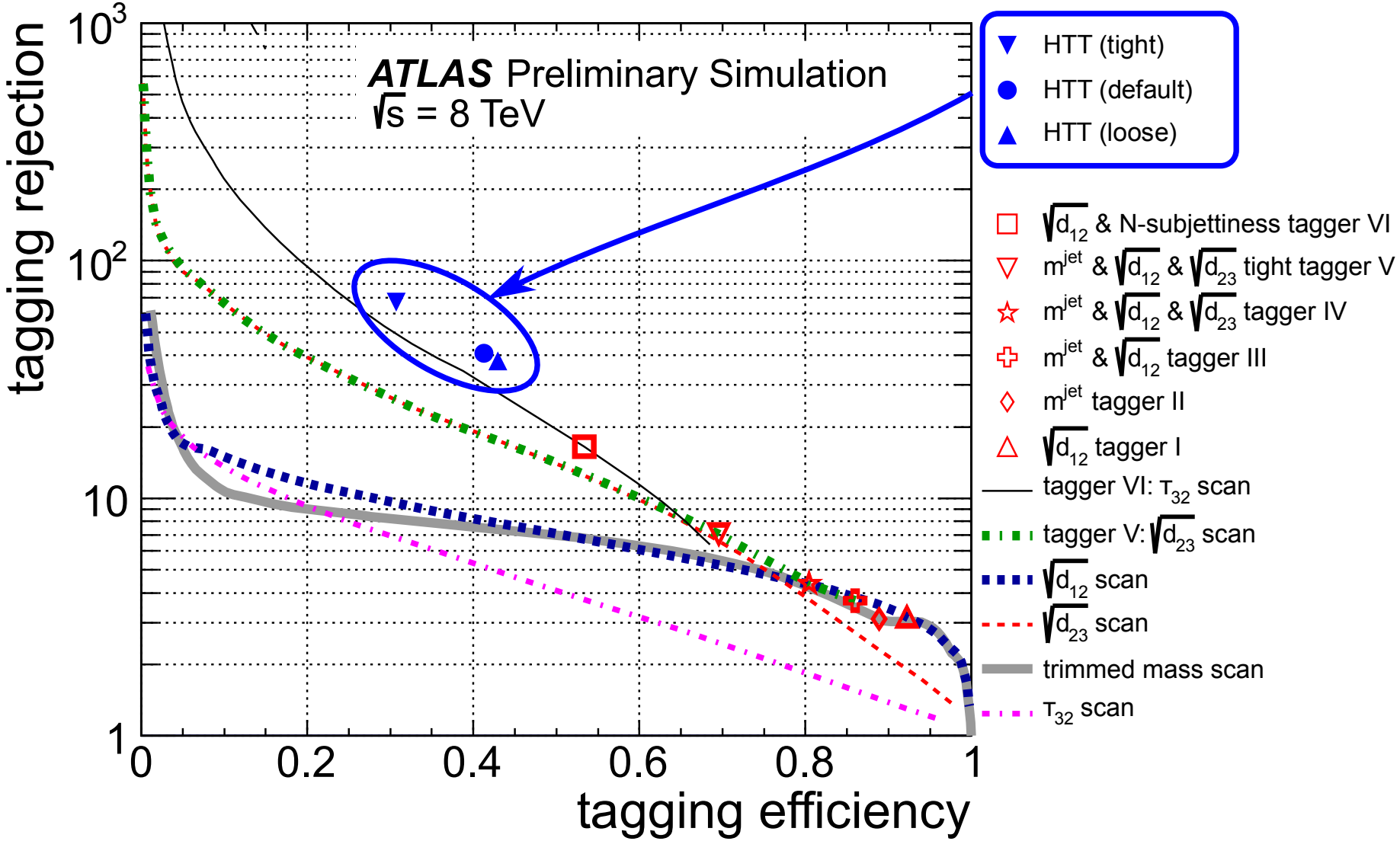
ATLAS-CONF-2014-003

Performance comparison in ATLAS: trimm. a-k_t R=1.0, CA R=1.2 (HEPTT), $p_{T,\text{jet}} > 550$ GeV



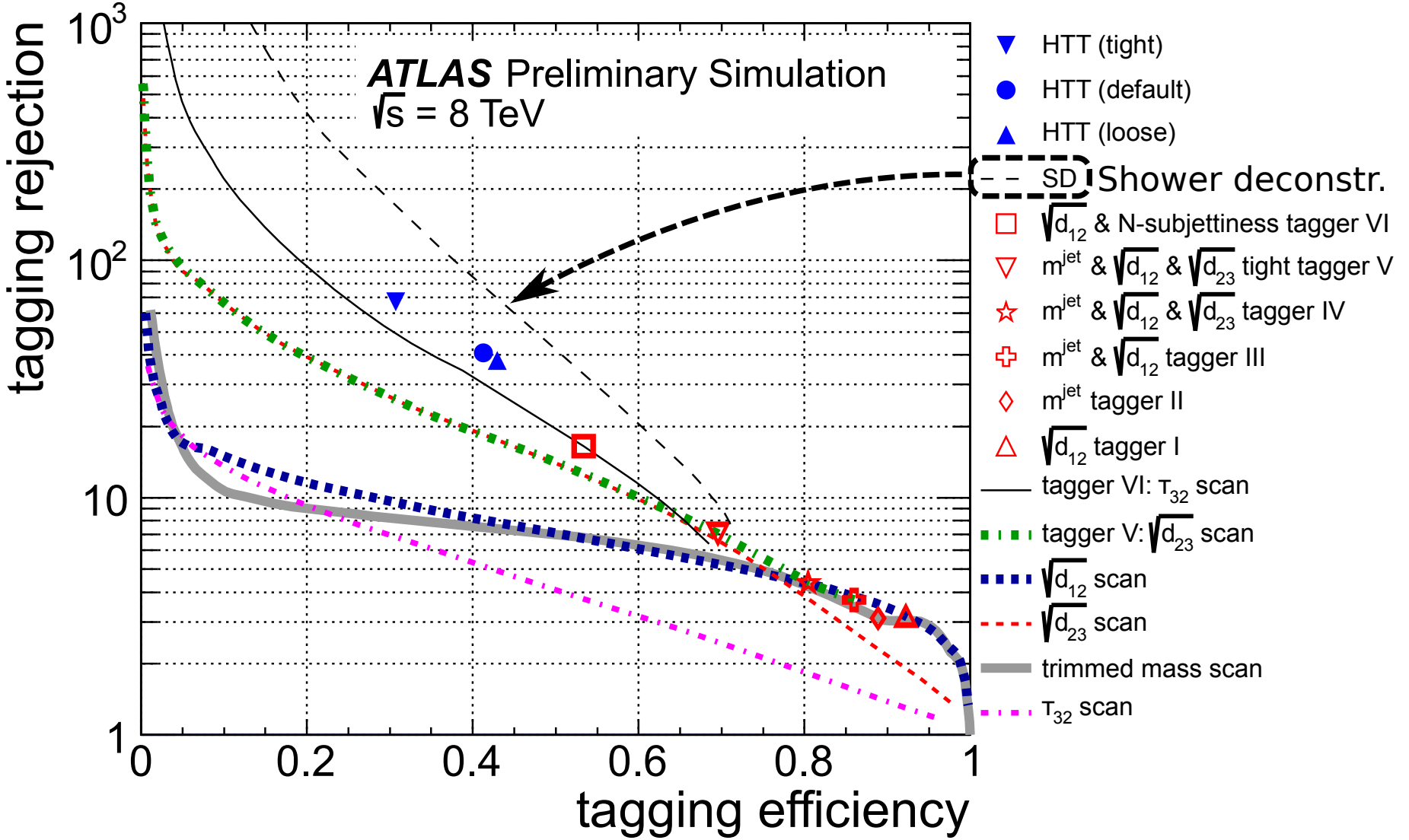
ATLAS-CONF-2014-003

Performance comparison in ATLAS: trimm. a-k_t R=1.0, CA R=1.2 (HEPTT), $p_{T,\text{jet}} > 550$ GeV



ATLAS-CONF-2014-003

Performance comparison in ATLAS: trimm. a-k_t R=1.0, CA R=1.2 (HEPTT), $p_{T,\text{jet}} > 550$ GeV



ATLAS-CONF-2014-003

Wrapping up

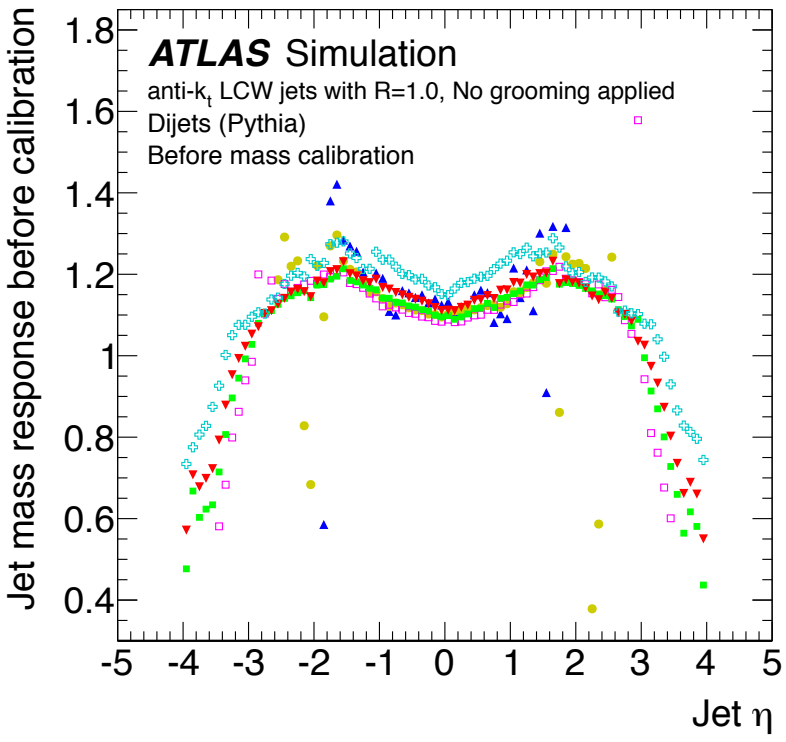
- very active field: many new theoretical and experimental developments every year
- widely used in searches for physics beyond the SM: see next talk from **Johannes Erdmann**
- getting ready for Run II: pileup safety, study systematics

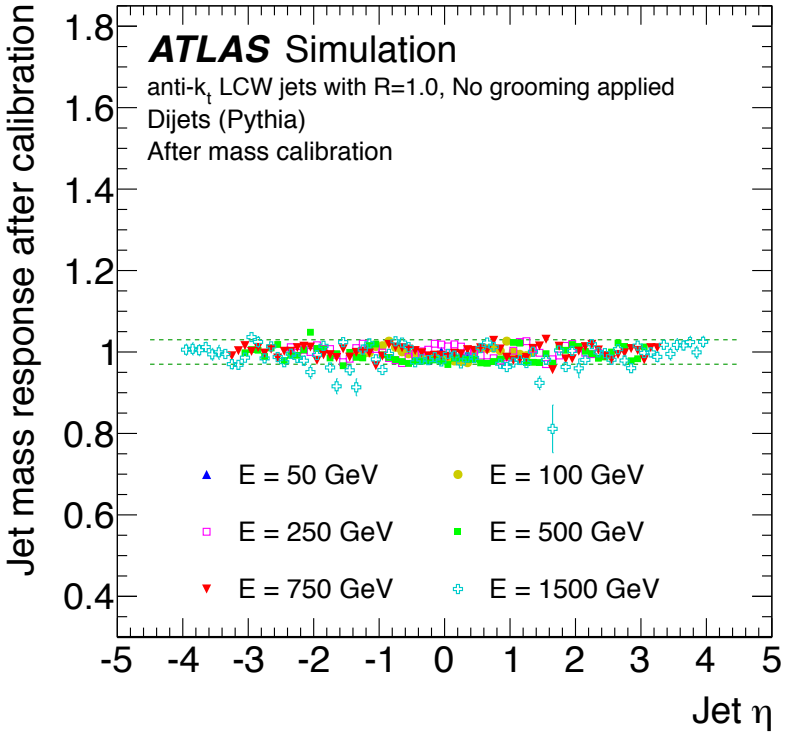
Thank you for the attention!

Backup

Large cone jet calibration (arXiv:1306.4945v1)

Large cone jets need specific calibration

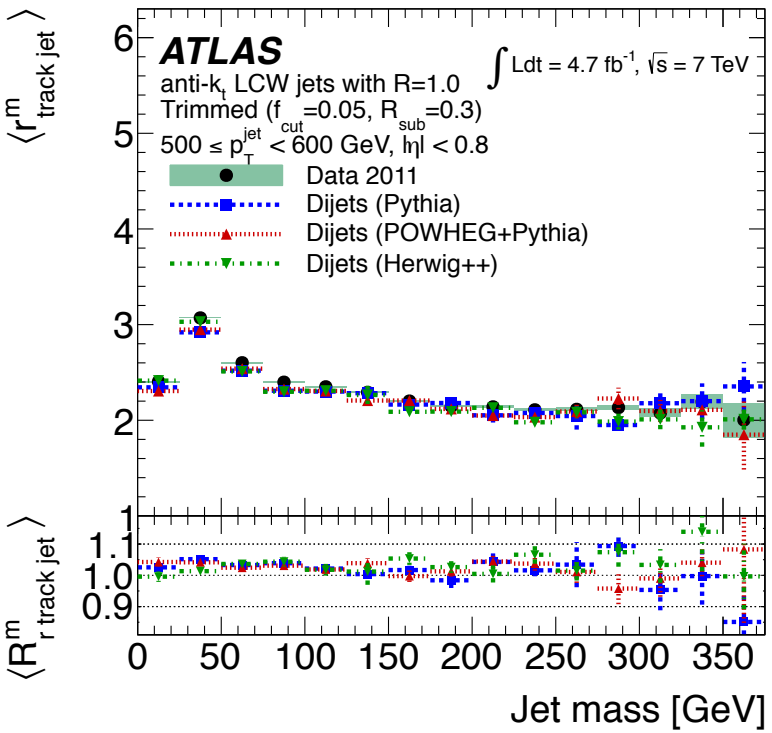


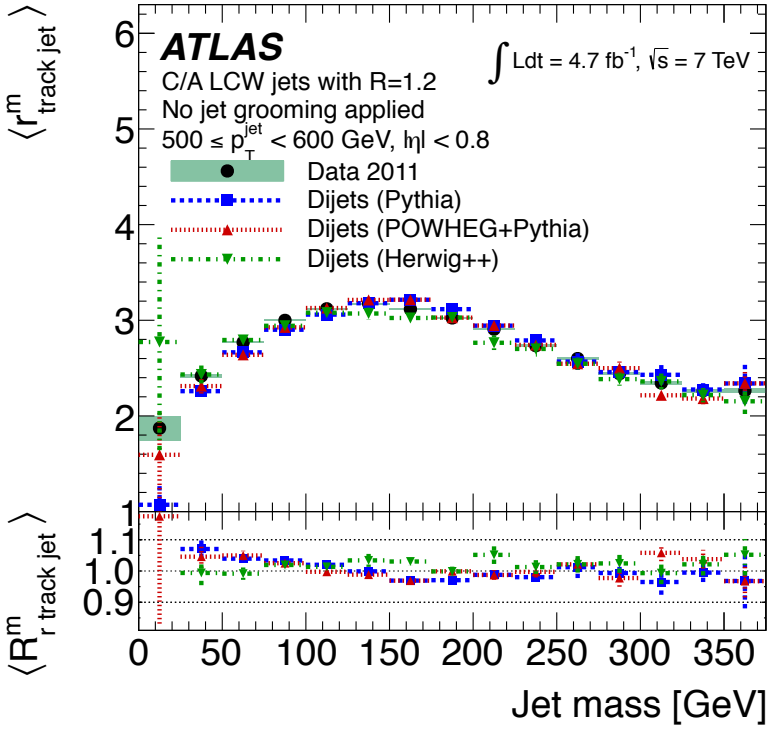


Large cone jet calibration (arXiv:1306.4945v1)

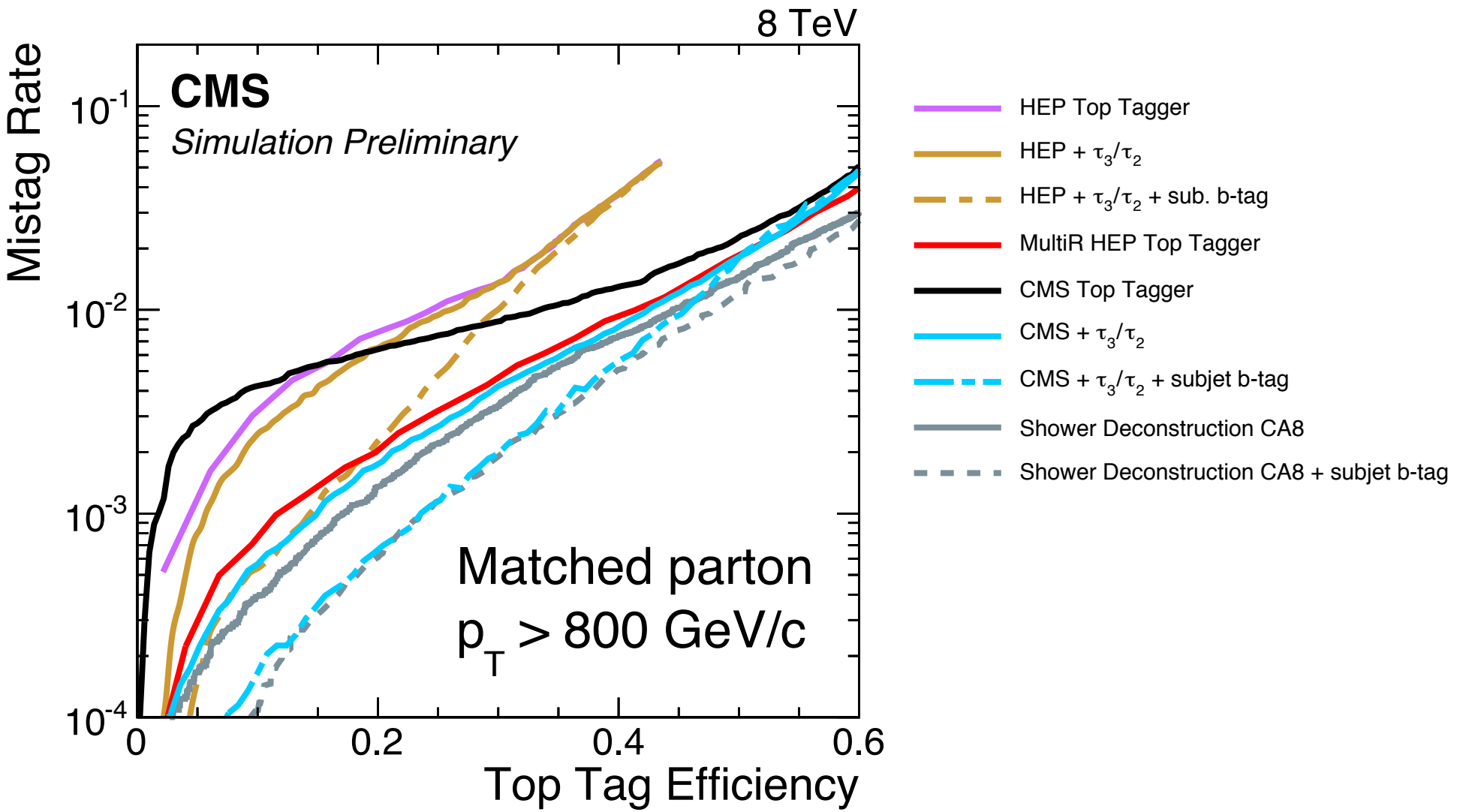
$$r^{m}_{\text{track jet}} = \frac{m^{\text{jet}}}{m^{\text{track jet}}}$$

$$R^{m}_{r\ \text{track jet}} = \frac{r^{m,\text{data}}_{\text{track jet}}}{r^{m,\text{MC}}_{\text{track jet}}}$$

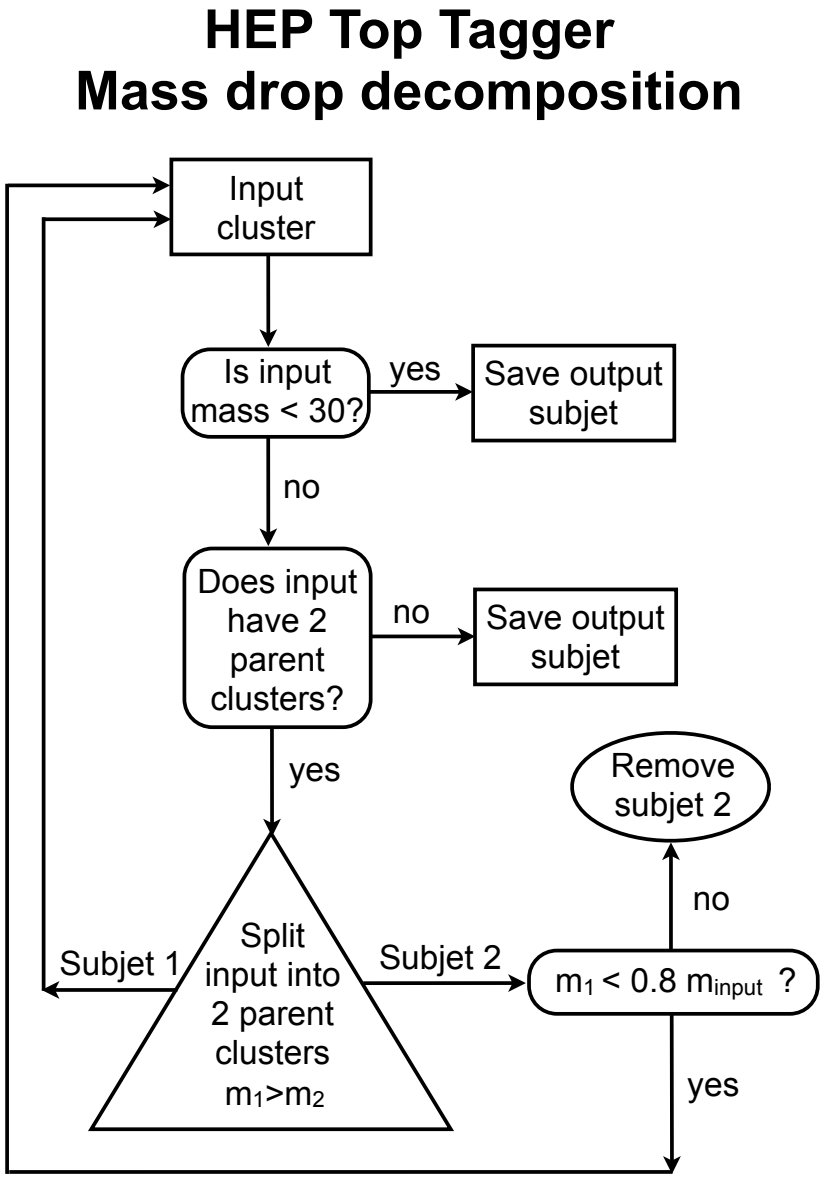




Performance comparison CMS



HEP tagger mass drop



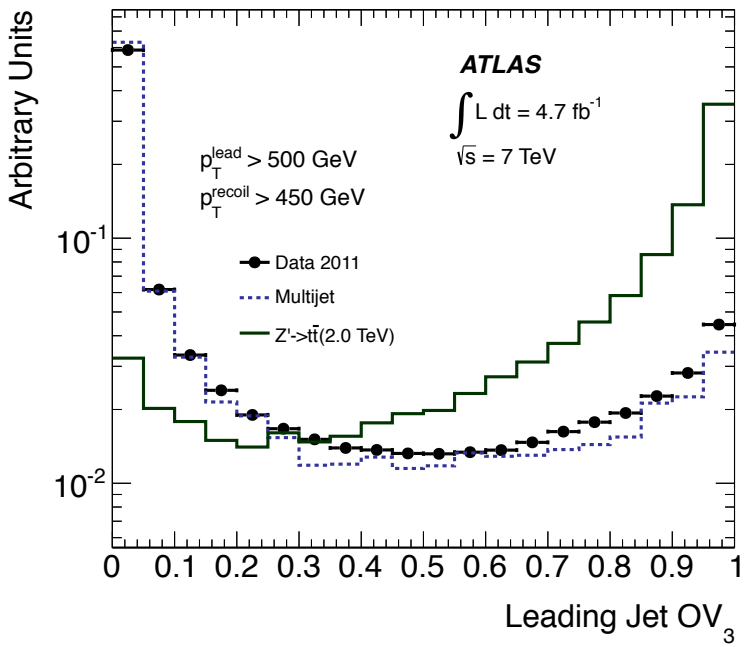
Other taggers in ATLAS

Reclustering:

- Consider the standard R=0.4 jets in an event
- use these jets as an input to recluster anti-k_t jets with larger R=0.8, 1.2
- use kinematic variables of the reclustered jets to cut

Template tagger:

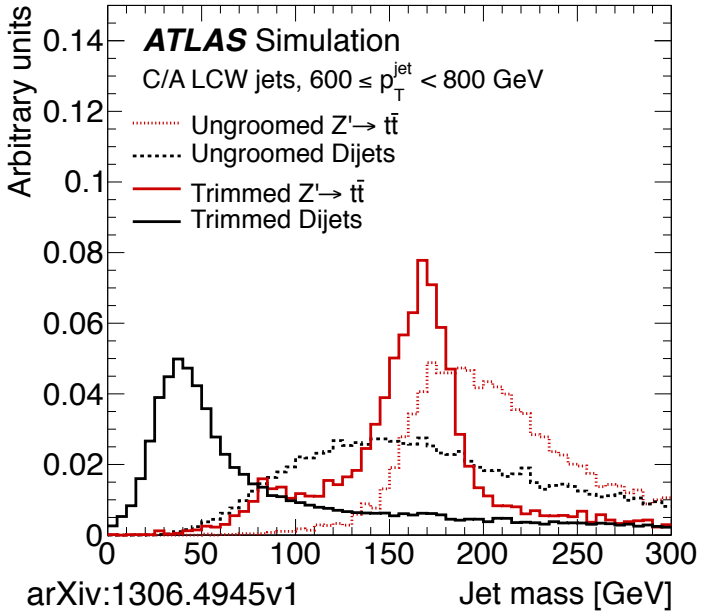
- Compare the energy flow for an event in data with the one from a big number of MC templates.
- from the coparisons compute a discriminator OV_3
- cut on OV_3 and on $|m - m_{\text{top}}|$



Jet grooming

Large radius jets:

- collect lots of soft QCD radiation
- suppress it to resolve hard decay product
- use jet grooming algorithms



Mass drop filtering

- isolate relatively symmetric subjets, with a significantly smaller mass than that of the original jet

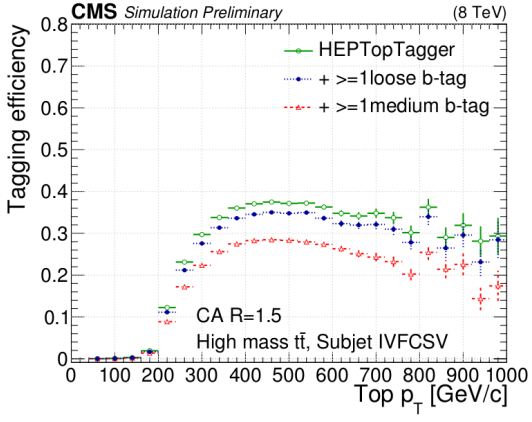
Trimming

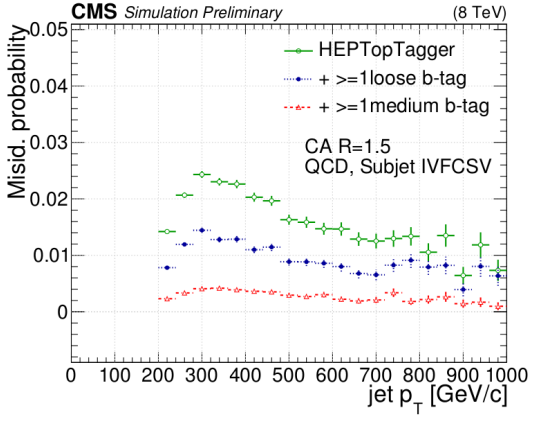
- reject softer subjets using a p_T fraction requirement ($p_{T,i}/p_{T,jet} < f$)

Pruning

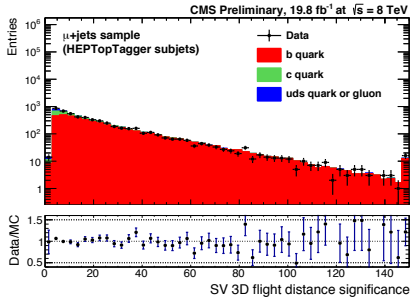
- reject soft and wide angle components in a jet ($p_{T,i}/p_{T,jet} > z_{cut}$ and $\Delta R_{ij} < R_{cut}(m_{jet}, p_{T,jet})$ at every recomb.)

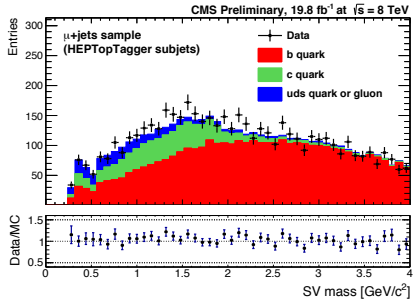
btag CMS

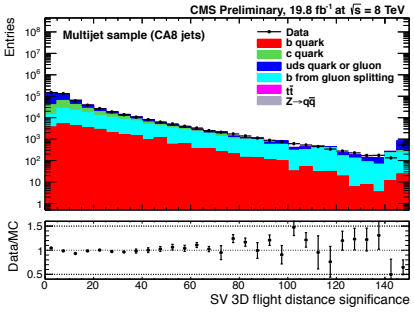


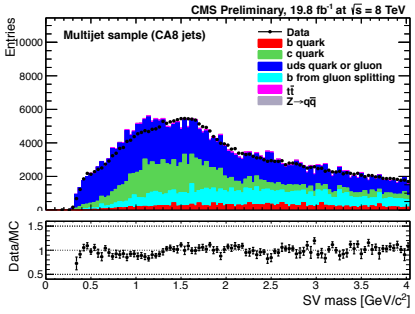


Tagging efficiency (left) and misidentification probability (right) as a function of the fat jet p_T. The HEPTopTagger performance (green) is combined with a subjet b-tagging requirement, for the medium (red) and loose (blue) operating points of the IVFCSV algorithm. Here, boosted top jets are defined as those jets that match within $\Delta R<1.5$ to a generator-level top that decays hadronically. Jets are reconstructed from charged-hadron-subtracted particle-flow candidates using the HEPTopTagger algorithm, based on the Cambridge/Aachen jet clustering algorithm with distance parameter R=1.5. A filtering procedure is applied to define exactly three subjets for each fat jet.









PUPPI

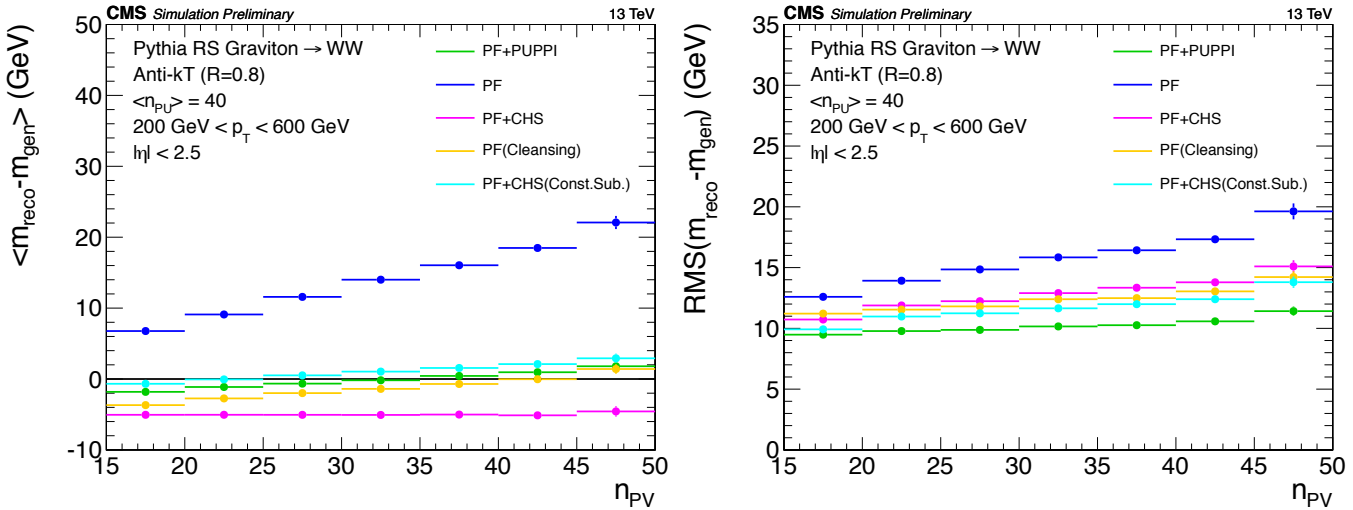


Figure 16: Mass response $< m_{\text{reco}} - m_{\text{gen}} >$ (left) and mass resolution quoted as RMS($m_{\text{reco}} - m_{\text{gen}}$) (right) for W jets as a function of the number reconstructed vertices.

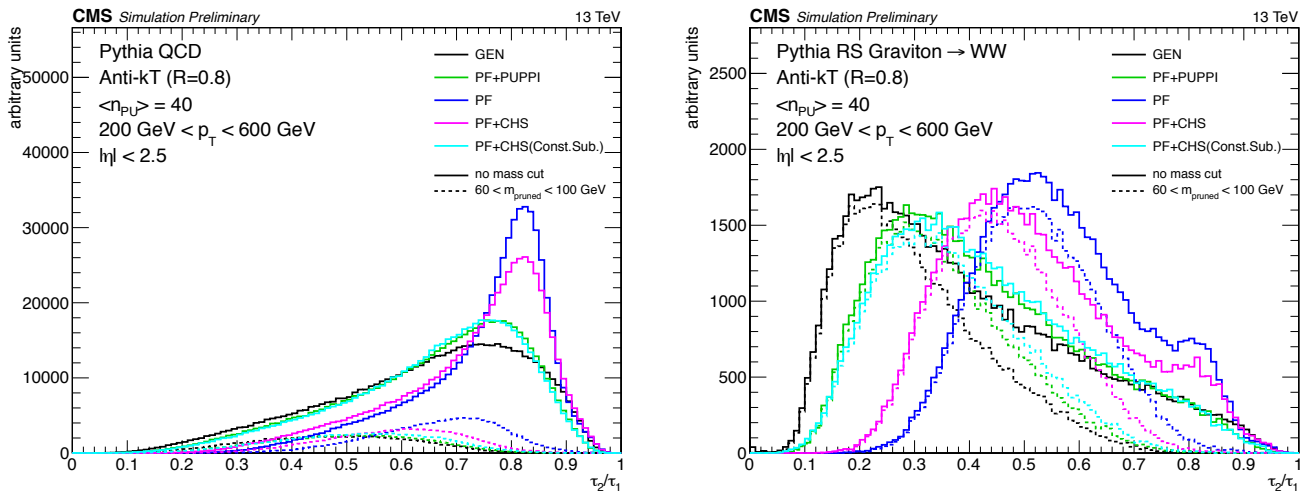


Figure 17: Leading jet N-subjettiness τ_2/τ_1 distribution: QCD jets (left) and W jets (right). The distribution is shown also after requiring the pruned mass to be in the range 60-100 GeV (dashed lines).

PUPPI 2

- General idea:
  - Define some local observable α that discriminates collinear vs soft diffuse structure in the neighborhood of a particle
  - Distribution of α for charged, central particles from PU is assumed to be representative of all PU (including forward)
  - Define particle weight based on α distribution of charged PU on event-by-event basis (so it works only for high PU...)
  - Rescale 4-momenta by these weights (before jet clustering)
- Authors tried some local observables and this one is best:

$$\alpha_i = \log \sum_{j \in \text{event}} \xi_{ij} \times \Theta(R_{\text{min}} \leq \Delta R_{ij} \leq R_0),$$

$$\text{where } \xi_{ij} = \frac{p_{Tj}}{\Delta R_{ij}}$$

These two parameters are detector dependent. Authors chose 0.02 and 0.3

from A. Giammanco

CMSTT SF

Sequential selection top-tagging efficiency - CMS Tagger, CMS Combined Tagger

$\|\eta\| < 1.0$				
Selection	**Data**	**MADGRAPH**	**POWHEG**	**MC@NLO**
$N_{\text{subjets}} \geq 3$	0.367±0.015	0.365±0.015	0.318±0.014	0.362±0.016
$m_{\text{min}} > 50$ (GeV/c^2)	0.719±0.023	0.754±0.022	0.735±0.024	0.725±0.024
$140 < m_{\text{jet}} < 250$ (GeV/c^2)	0.954±0.012	0.928±0.015	0.917±0.017	0.928±0.016
$\tau_3/\tau_2 < 0.55$	0.554±0.030	0.573±0.030	0.559±0.032	0.587±0.033
subjet b-tag CSV-medium	0.658±0.039	0.704±0.037	0.718±0.039	0.687±0.040

$1.0 < \|\eta\| < 2.4$				
Selection	**Data**	**MADGRAPH**	**POWHEG**	**MC@NLO**
$N_{\text{subjets}} \geq 3$	0.326±0.022	0.323±0.022	0.314±0.021	0.291±0.019
$m_{\text{min}} > 50$ (GeV/c^2)	0.456±0.041	0.661±0.040	0.619±0.040	0.615±0.037
$140 < m_{\text{jet}} < 250$ (GeV/c^2)	0.866±0.042	0.936±0.025	0.939±0.025	0.936±0.024
$\tau_3/\tau_2 < 0.55$	0.362±0.063	0.453±0.053	0.428±0.053	0.447±0.051
subjet b-tag CSV-medium	0.905±0.064	0.680±0.074	0.595±0.080	0.702±0.070

Data-simulation scale factor for sequential selections - CMS Tagger, CMS Combined Tagger

$\|\eta\| < 1.0$			
Selection	**MADGRAPH**	**POWHEG**	**MC@NLO**
$N_{\text{subjets}} \geq 3$	1.006 ± 0.057	1.153 ± 0.069	1.014 ± 0.060
$m_{\text{min}} > 50$ GeV/c^2	0.954 ± 0.041	0.978 ± 0.044	0.992 ± 0.046
140 GeV/$c^2 < m_{\text{jet}} < 250$ GeV/c^2	1.028 ± 0.021	1.040 ± 0.024	1.028 ± 0.023
$\tau_3/\tau_2 < 0.55$	0.967 ± 0.073	0.990 ± 0.079	0.943 ± 0.073
subjet b-tag CSV-medium	0.935 ± 0.074	0.915 ± 0.074	0.957 ± 0.079

$1.0 < \|\eta\| < 2.4$			
Selection	**MADGRAPH**	**POWHEG**	**MC@NLO**
$N_{\text{subjets}} \geq 3$	1.010 ± 0.097	1.037 ± 0.099	1.122 ± 0.106
$m_{\text{min}} > 50$ GeV/c^2	0.689 ± 0.075	0.737 ± 0.081	0.741 ± 0.081
140 GeV/$c^2 < m_{\text{jet}} < 250$ GeV/c^2	0.925 ± 0.051	0.922 ± 0.051	0.925 ± 0.051
$\tau_3/\tau_2 < 0.55$	0.800 ± 0.168	0.845 ± 0.181	0.810 ± 0.169
subjet b-tag CSV-medium	1.331 ± 0.172	1.52 ± 0.232	1.289 ± 0.157

HEPTT SF

Cumulative data-simulation scale factor - HEP Top Tagger, HEP Combined Tagger

$\|\eta\| < 1.0$				
Tagger	**p_T bin (GeV/c)**	**MADGRAPH**	**POWHEG**	**MC@NLO**
HEP Combined WP2	$200 < p_T < 250$	0.91 ± 0.04	0.92 ± 0.04	0.88 ± 0.04
	$250 < p_T < 400$	0.93 ± 0.03	0.95 ± 0.03	0.93 ± 0.03
	$p_T > 400$	1.15 ± 0.07	1.36 ± 0.07	1.19 ± 0.07
HEP Combined WP3	$200 < p_T < 250$	0.86 ± 0.05	0.86 ± 0.05	0.86 ± 0.05
	$250 < p_T < 400$	0.91 ± 0.04	0.93 ± 0.04	0.93 ± 0.04
	$p_T > 400$	0.98 ± 0.09	1.10 ± 0.12	1.10 ± 0.12

$1.0 < \|\eta\| < 2.4$				
Tagger	**p_T bin (GeV/c)**	**MADGRAPH**	**POWHEG**	**MC@NLO**
HEP Combined WP2	$200 < p_T < 250$	0.95 ± 0.05	0.93 ± 0.06	0.93 ± 0.05
	$250 < p_T < 400$	0.91 ± 0.04	0.95 ± 0.05	0.95 ± 0.04
	$p_T > 400$	0.85 ± 0.11	0.95 ± 0.15	0.99 ± 0.13
HEP Combined WP3	$200 < p_T < 250$	1.02 ± 0.07	1.00 ± 0.08	0.96 ± 0.07
	$250 < p_T < 400$	0.90 ± 0.05	0.97 ± 0.06	0.93 ± 0.05
	$p_T > 400$	0.85 ± 0.16	1.00 ± 0.22	0.99 ± 0.19

Sequential data-simulation scale factor - HEP Combined Tagger WP3

$\|\eta\| < 1.0$				
Tagger	**p_T bin (GeV/c)**	**MADGRAPH**	**POWHEG**	**MC@NLO**
HEP top mass selection	$200 < p_T < 250$	0.92 ± 0.02	0.94 ± 0.03	0.92 ± 0.02
	$250 < p_T < 400$	0.93 ± 0.02	0.96 ± 0.02	0.94 ± 0.02
	$p_T > 400$	1.03 ± 0.03	1.07 ± 0.04	1.04 ± 0.04
HEP W mass selection	$200 < p_T < 250$	0.98 ± 0.03	0.98 ± 0.04	0.96 ± 0.03
	$250 < p_T < 400$	0.99 ± 0.02	0.99 ± 0.03	0.99 ± 0.02
	$p_T > 400$	1.11 ± 0.05	1.27 ± 0.07	1.14 ± 0.06
N-subjettiness selection	$200 < p_T < 250$	0.95 ± 0.04	0.93 ± 0.04	0.90 ± 0.03
	$250 < p_T < 400$	0.98 ± 0.03	0.98 ± 0.03	0.94 ± 0.03
	$p_T > 400$	0.85 ± 0.06	0.81 ± 0.07	0.84 ± 0.06

$1.0 < \|\eta\| < 2.4$				
Tagger	**p_T bin (GeV/c)**	**MADGRAPH**	**POWHEG**	**MC@NLO**
HEP top mass selection	$200 < p_T < 250$	0.89 ± 0.03	0.89 ± 0.04	0.90 ± 0.03
	$250 < p_T < 400$	0.92 ± 0.02	0.96 ± 0.03	0.97 ± 0.02
	$p_T > 400$	0.92 ± 0.07	0.98 ± 0.08	1.07 ± 0.08
HEP W mass selection	$200 < p_T < 250$	1.07 ± 0.04	1.04 ± 0.05	1.03 ± 0.04
	$250 < p_T < 400$	0.99 ± 0.03	0.99 ± 0.04	0.98 ± 0.03
	$p_T > 400$	0.92 ± 0.10	0.97 ± 0.13	0.92 ± 0.11
N-subjettiness selection	$200 < p_T < 250$	1.07 ± 0.05	1.07 ± 0.05	1.03 ± 0.05
	$250 < p_T < 400$	0.99 ± 0.04	1.03 ± 0.05	0.98 ± 0.04
	$p_T > 400$	1.01 ± 0.13	1.06 ± 0.17	1.00 ± 0.14

Wtag

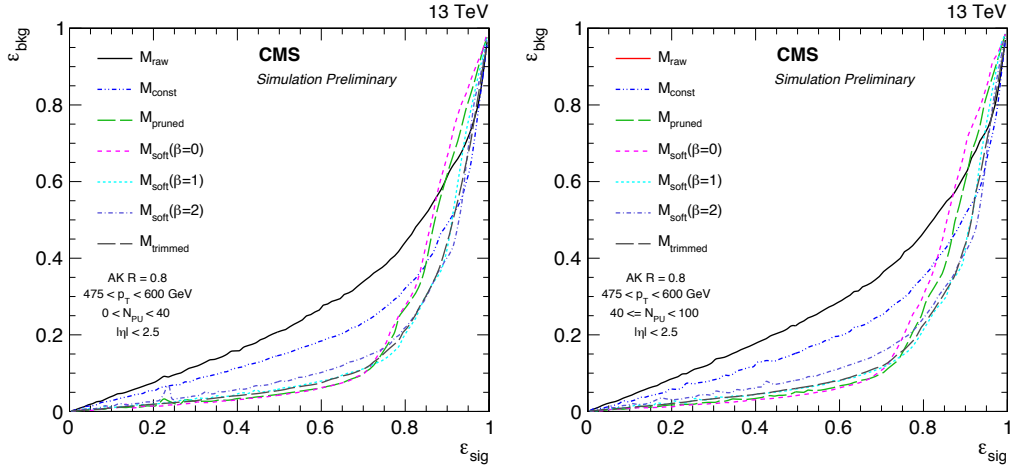


Figure 16: Comparison of various discriminant mass observable performance in the 475-600 GeV jet p_T bin, obtained considering simulated RS Graviton decaying in WW final state as source of W jet. Left: a comparison in the low pileup region $N_{PU} < 40$, and right: a comparison in the high pileup region $40 \leq N_{PU}$.

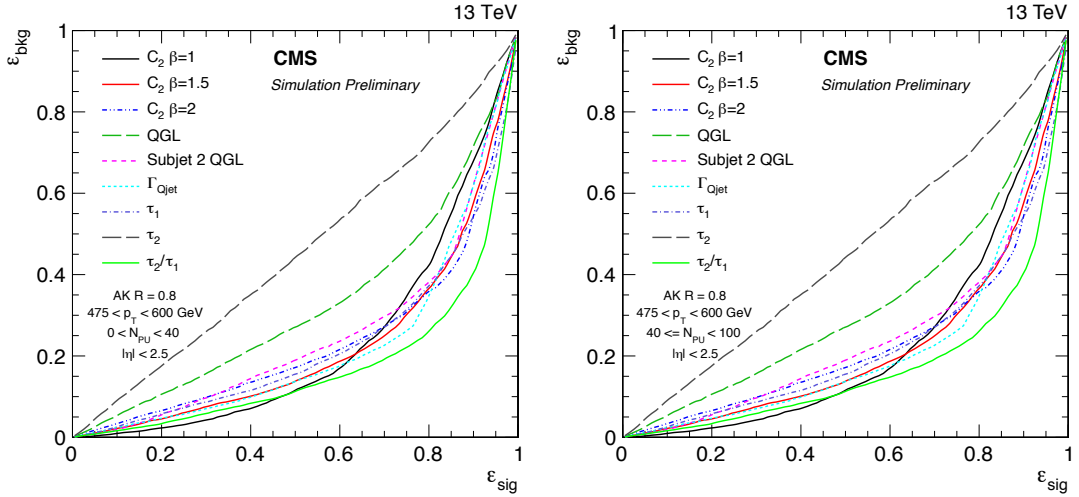
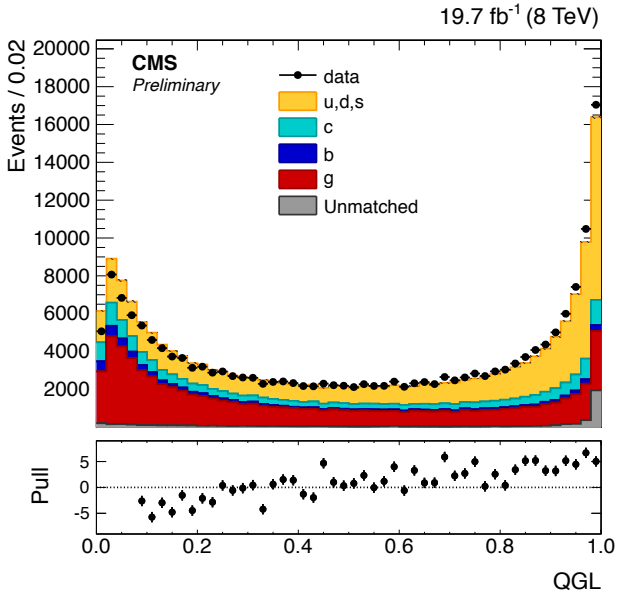
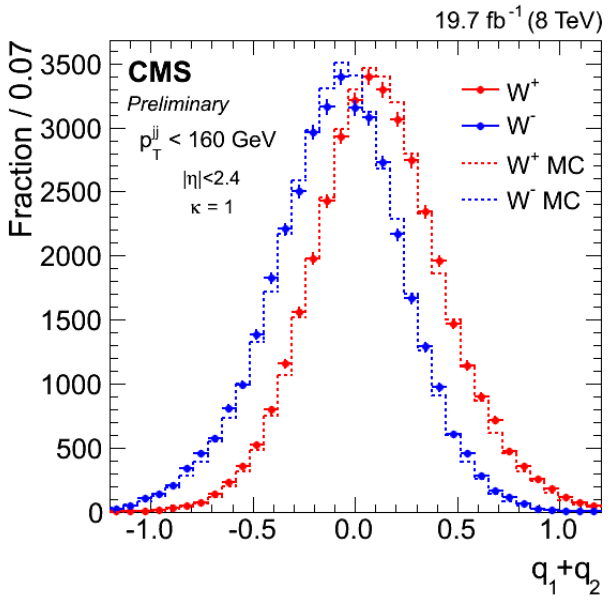


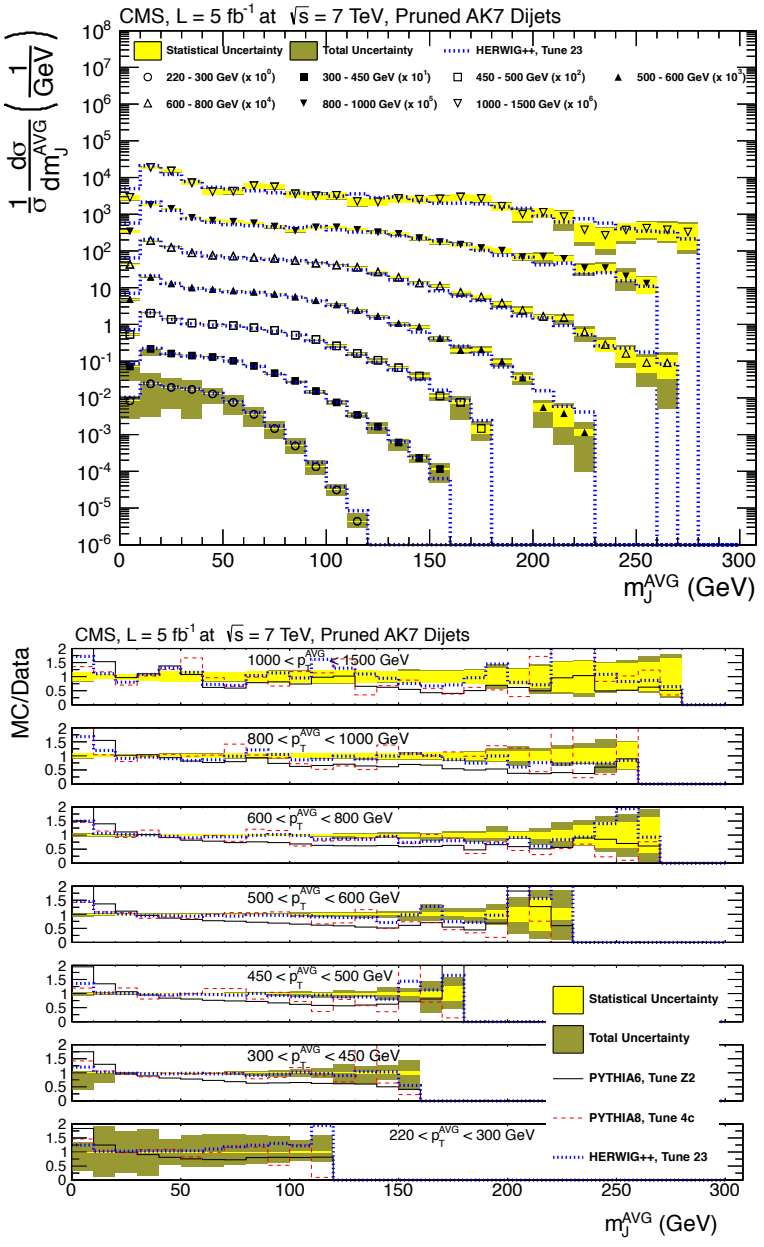
Figure 17: Comparison of various substructure observable performance in the jet p_T bin, 475-600 GeV, considering W jet in RS $G \rightarrow$ WW ($M_G = 1$ TeV) events as signal. (Left) Comparison in the low pileup region $0 < N_{PU} < 40$, (Right) high pileup one $40 \leq N_{PU} < 100$.

Wtag 2

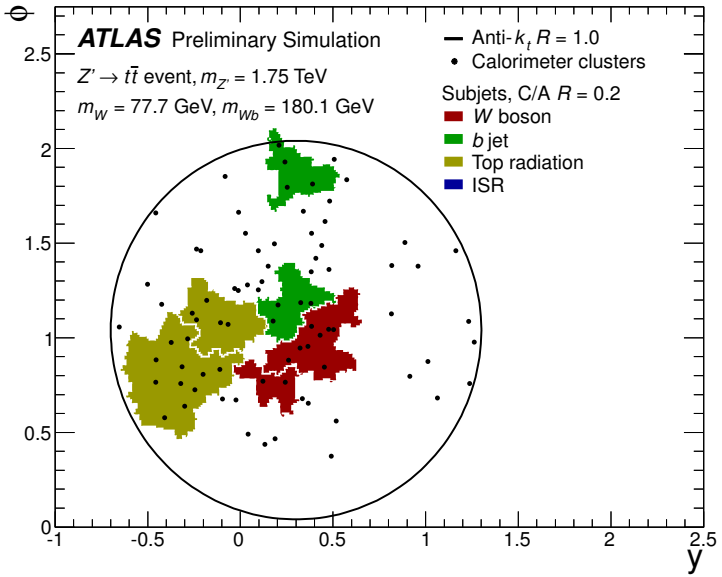

19.7 fb-1 (8 TeV)
CMS
Preliminary
data
u,d,s
c
b
g
Unmatched
Events / 0.02
Pull
QGL

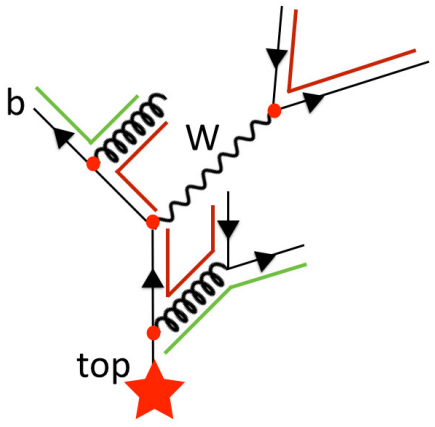

19.7 fb-1 (8 TeV)
CMS
Preliminary
pTjj < 160 GeV
|η|<2.4
κ = 1
W+
W-
W+ MC
W- MC
Fraction / 0.07
q1+q2

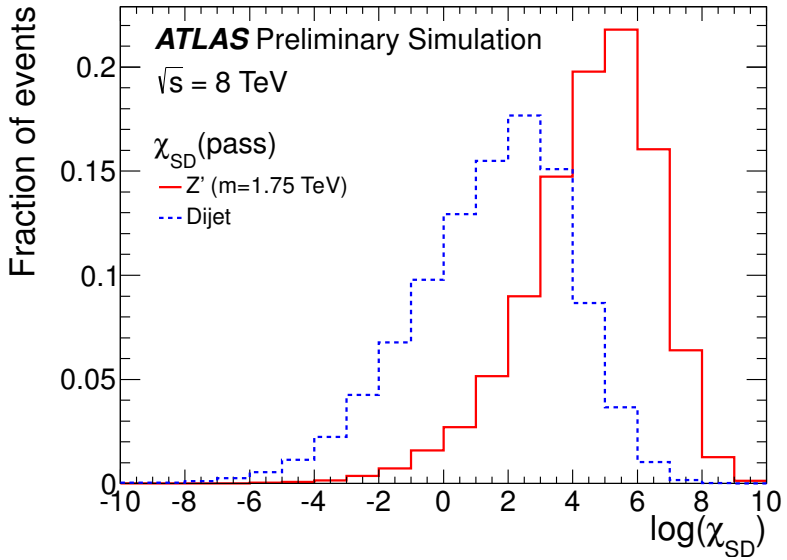
grooming CMS



Shower deconstruction





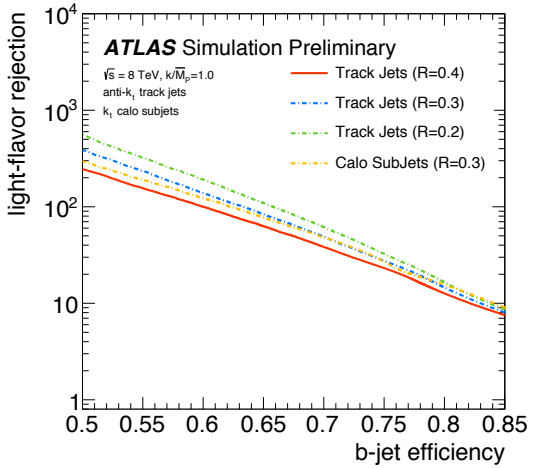


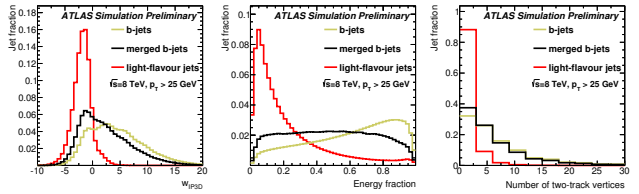
IVF

1. Start with all tracks, regardless of jets
2. Look for **Seed Tracks** with large **Impact Parameter** (distance to primary vertex)
3. Look for **track clusters** around Seed Tracks
4. Fit the position of the secondary vertex from track clusters
5. Apply some cleaning steps (e.g. merge vertices that are too close together)

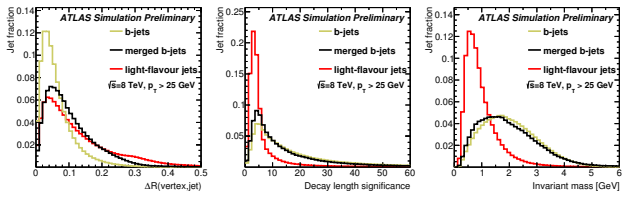
from D. Nowatschin

Btag ATLAS

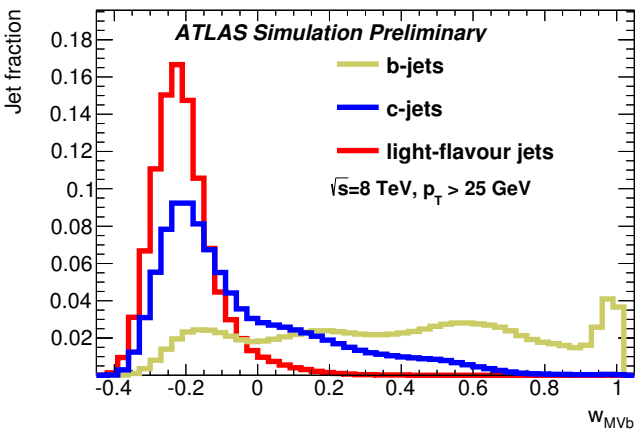


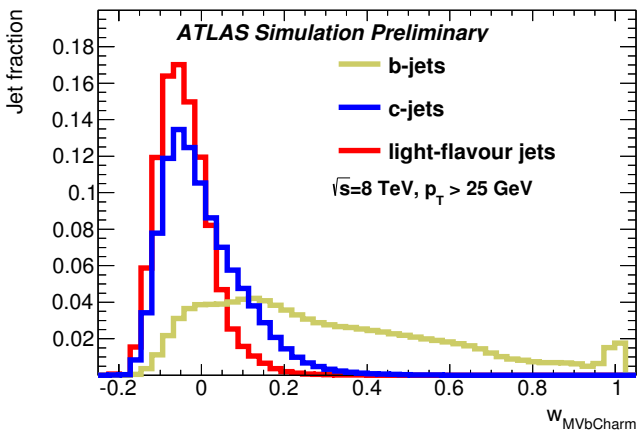


(a) IP3D weight (b) energy fraction (JetFitter based) (c) number of two-track vertices (JetFitter based)

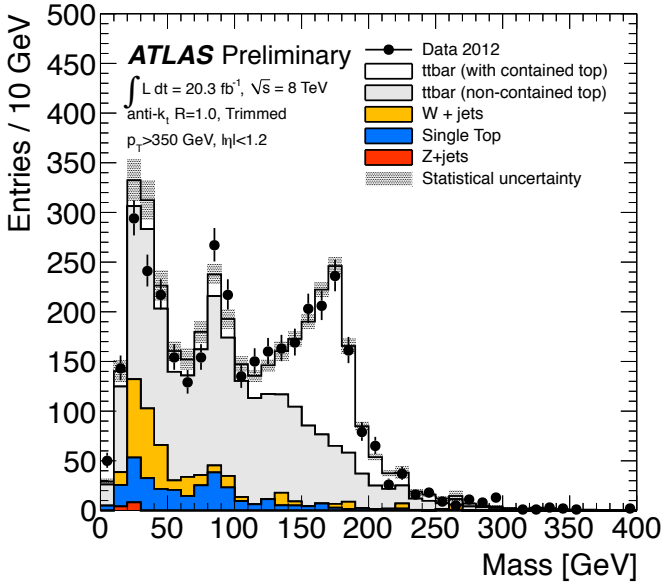


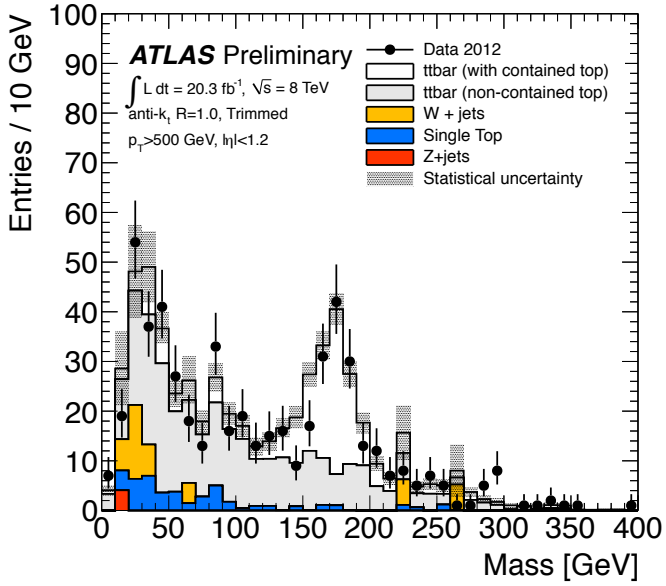
(d) ΔR(vertex, jet) (SV1 based) (e) decay length significance (SV1 based) (f) invariant mass (SV1 based)

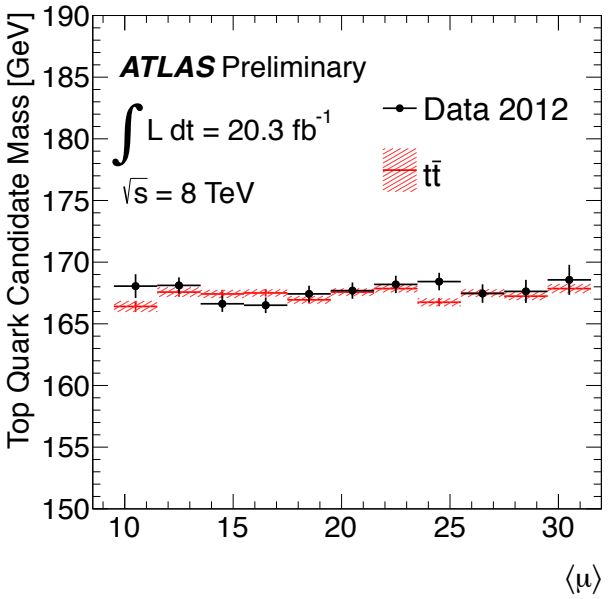


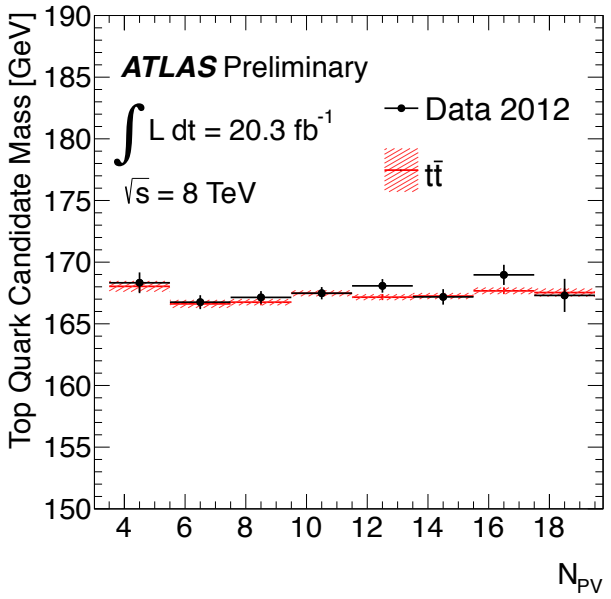


Top Tag

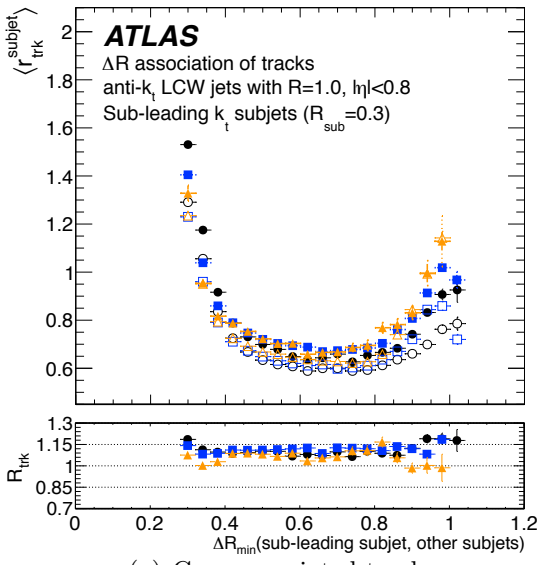

ATLAS Preliminary
∫ L dt = 20.3 fb-1, √s = 8 TeV
anti-kt R=1.0, Trimmed
pT>350 GeV, |η|<1.2
Data 2012
ttbar (with contained top)
ttbar (non-contained top)
W + jets
Single Top
Z+jets
Statistical uncertainty
Entries / 10 GeV
Mass [GeV]


ATLAS Preliminary
∫ L dt = 20.3 fb-1, √s = 8 TeV
anti-kt R=1.0, Trimmed
pT>500 GeV, |η|<1.2
Data 2012
ttbar (with contained top)
ttbar (non-contained top)
W + jets
Single Top
Z+jets
Statistical uncertainty
Entries / 10 GeV
Mass [GeV]

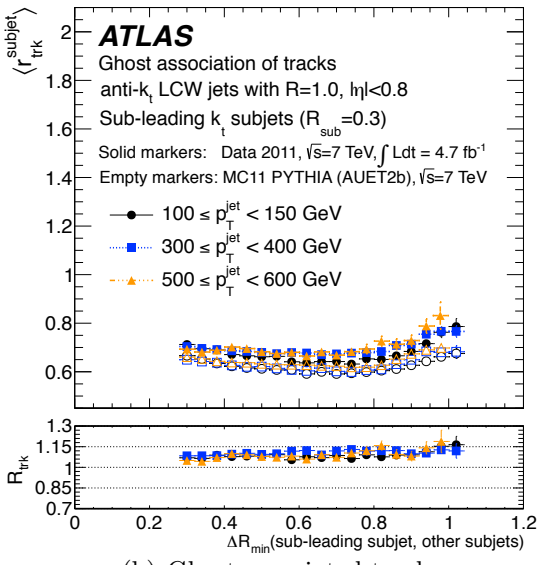

ATLAS Preliminary
∫ L dt = 20.3 fb-1
√s = 8 TeV
Data 2012
tt̄
Top Quark Candidate Mass [GeV]
⟨μ⟩


ATLAS Preliminary
∫ L dt = 20.3 fb-1
√s = 8 TeV
Data 2012
tt̄
Top Quark Candidate Mass [GeV]
NPV

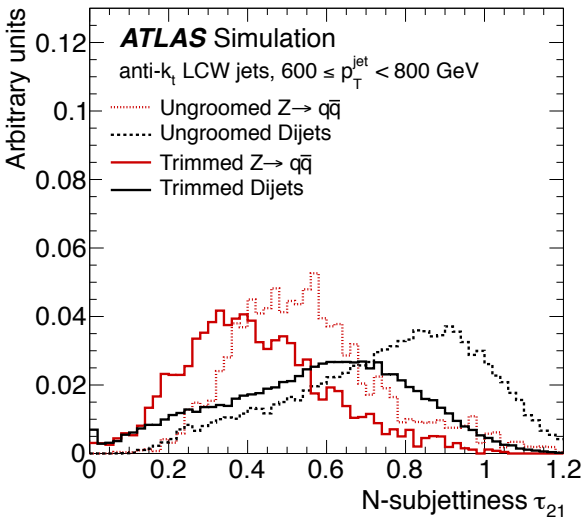
Jets



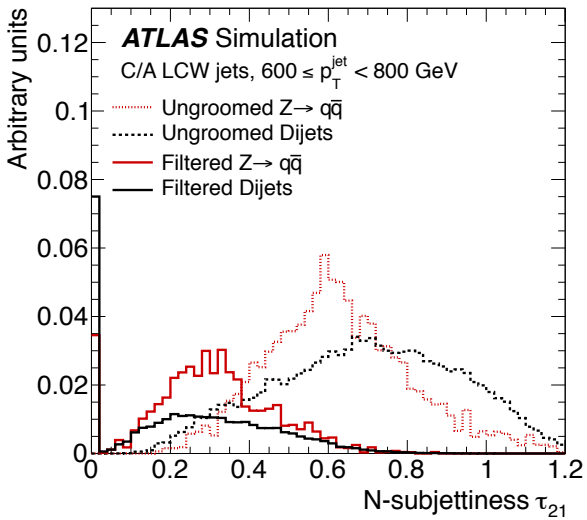
(a) Cone associated tracks



(b) Ghost associated tracks

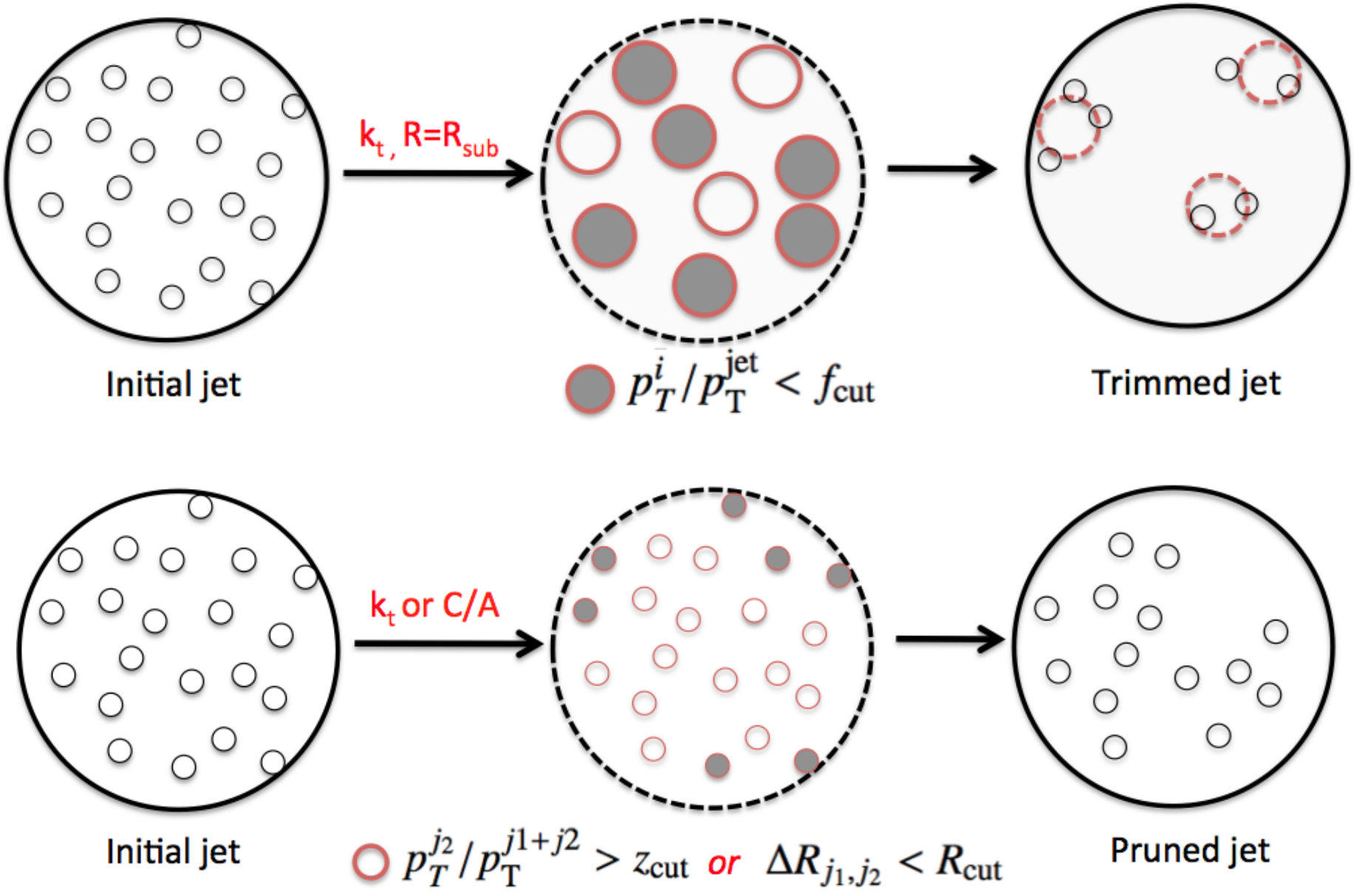


(a) anti-k_t, $R = 1.0$

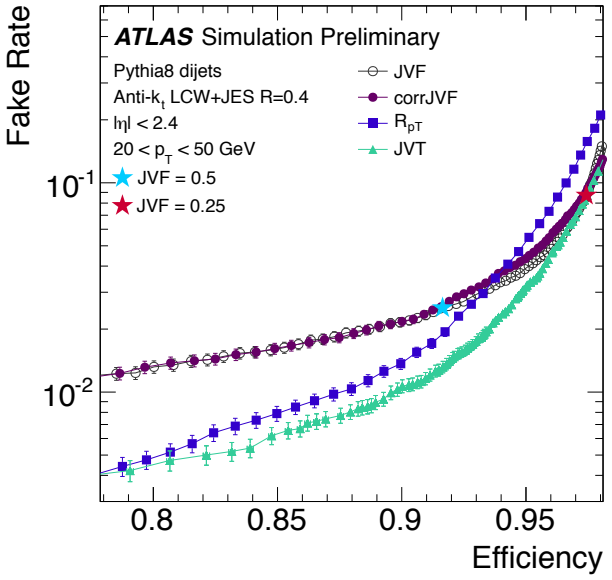


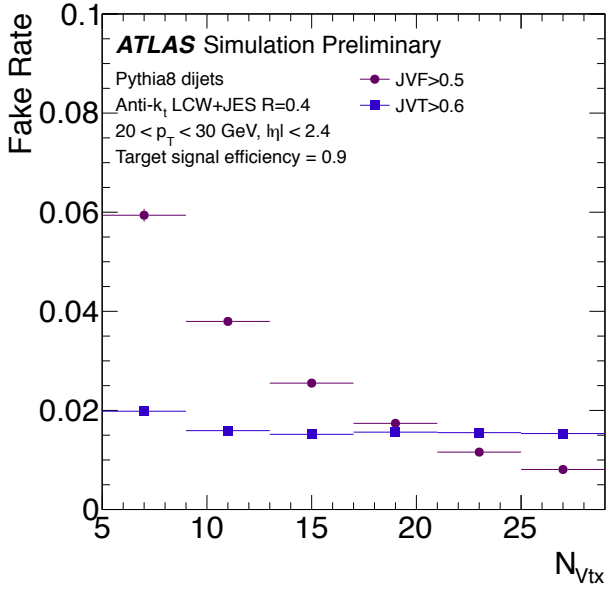
(b) C/A, $R = 1.2$

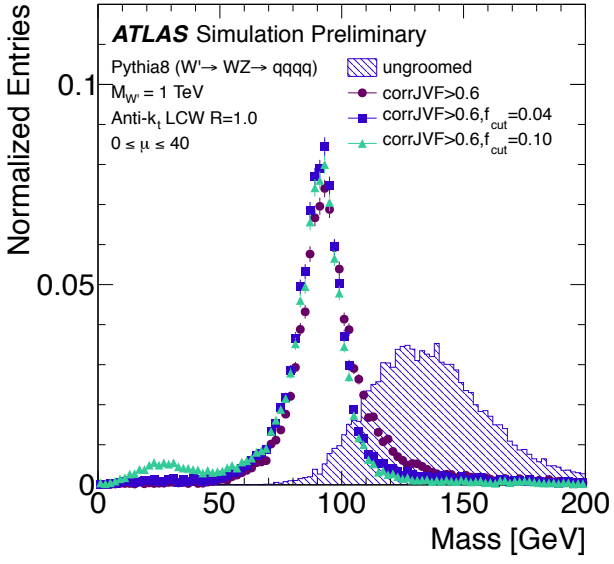
Grooming

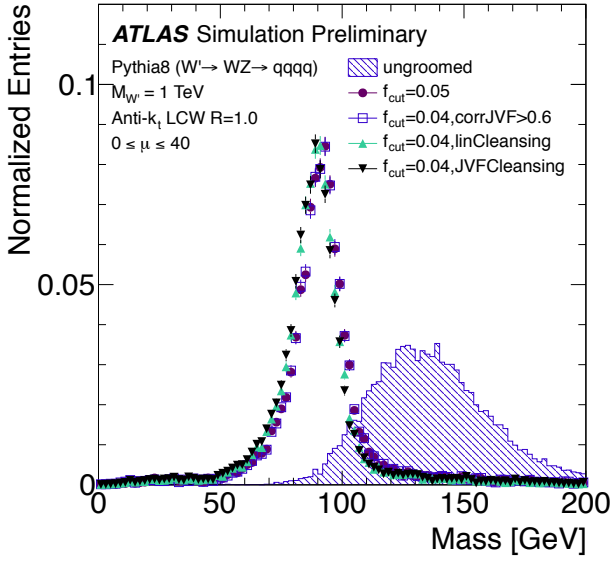

k_t , R=R_{sub}
Initial jet
$p_T^i / p_T^{\text{jet}} < f_{\text{cut}}$
Trimmed jet
k_t or C/A
Initial jet
$p_T^{j2} / p_T^{j1+j2} > z_{\text{cut}}$ or $\Delta R_{j_1,j_2} < R_{\text{cut}}$
Pruned jet

PUID and cleansing

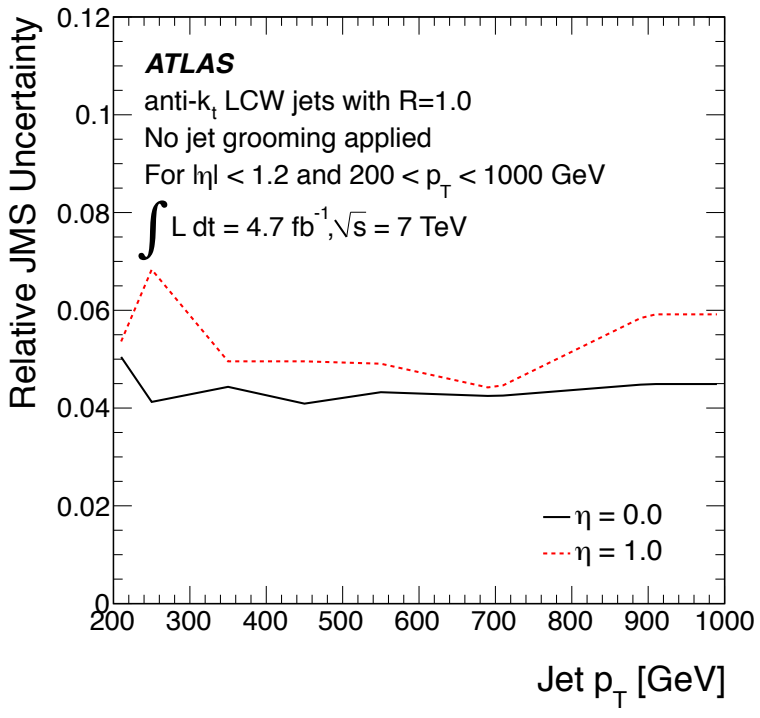




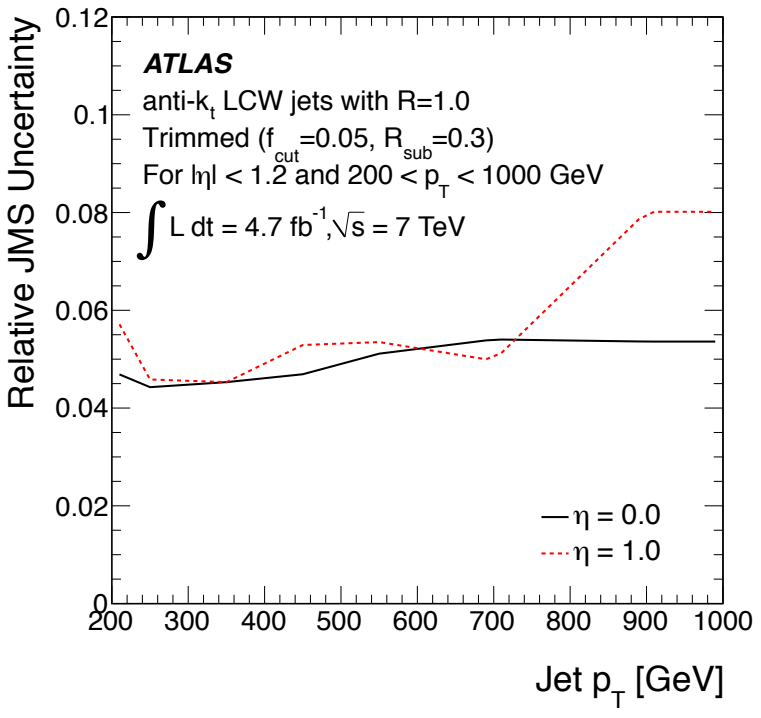




JMS relative systematic uncertainties

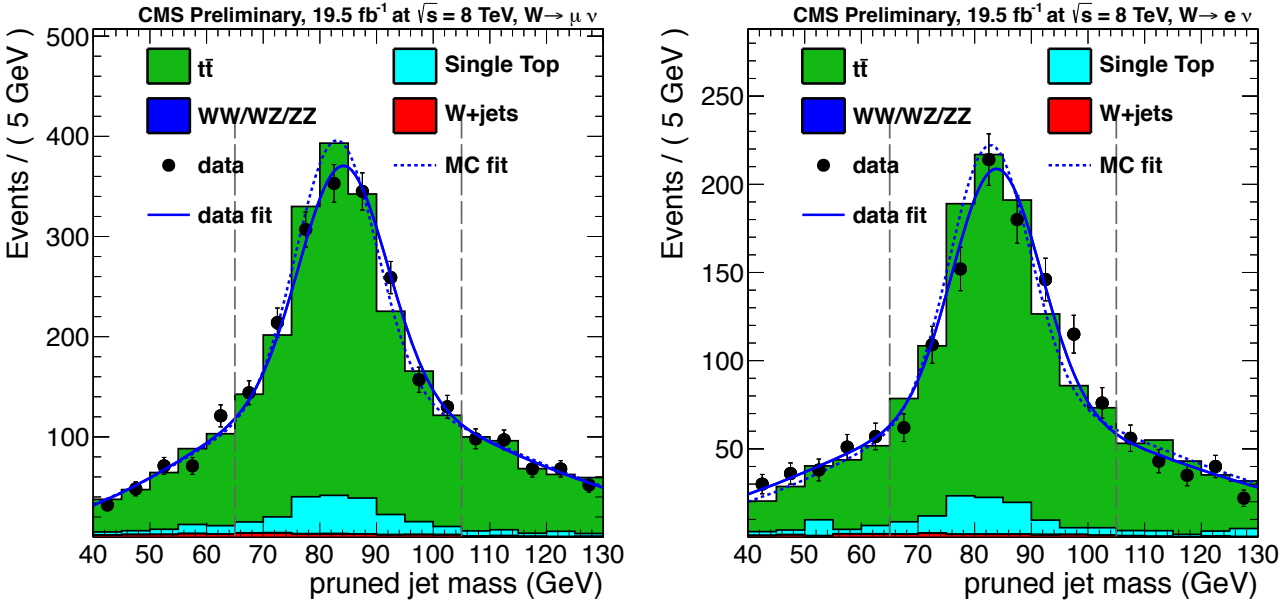


(a) anti-k_t, $R = 1.0$ (no jet grooming)



(b) anti-k_t, $R = 1.0$ (trimming)

W-jet mass calibration



$$\langle m \rangle_{\text{sim}} = 83.4 \pm 0.4\ \text{GeV}\ ,\ \sigma_{\text{sim}} = 7.5 \pm 0.4\ \text{GeV} \quad (5)$$

$$\langle m \rangle_{\text{data}} = 84.5 \pm 0.4\ \text{GeV}\ ,\ \sigma_{\text{data}} = 8.7 \pm 0.6\ \text{GeV} \quad (6)$$

We find that both the W-jet mass scale and resolution in data are larger than that in simulation. We should shift the simulation $\langle m \rangle$ by $+1.1 \pm 0.6$ GeV and enlarge the simulation σ by $16\% \pm 9\%$ to correct for the difference in data/simulation.

W-jet mass calibration

Source	Effect on the scale factor
Parton showering	6.0%
Pileup	1.8%
Jet mass scale	$< 0.5\%$
Jet mass resolution	0.7%
Jet energy scale	1.9%
Jet energy resolution	0.9%
Lepton ID	$< 0.5\%$
b-tagging	$< 0.5\%$
MET	$< 0.5\%$
Total	6.7%