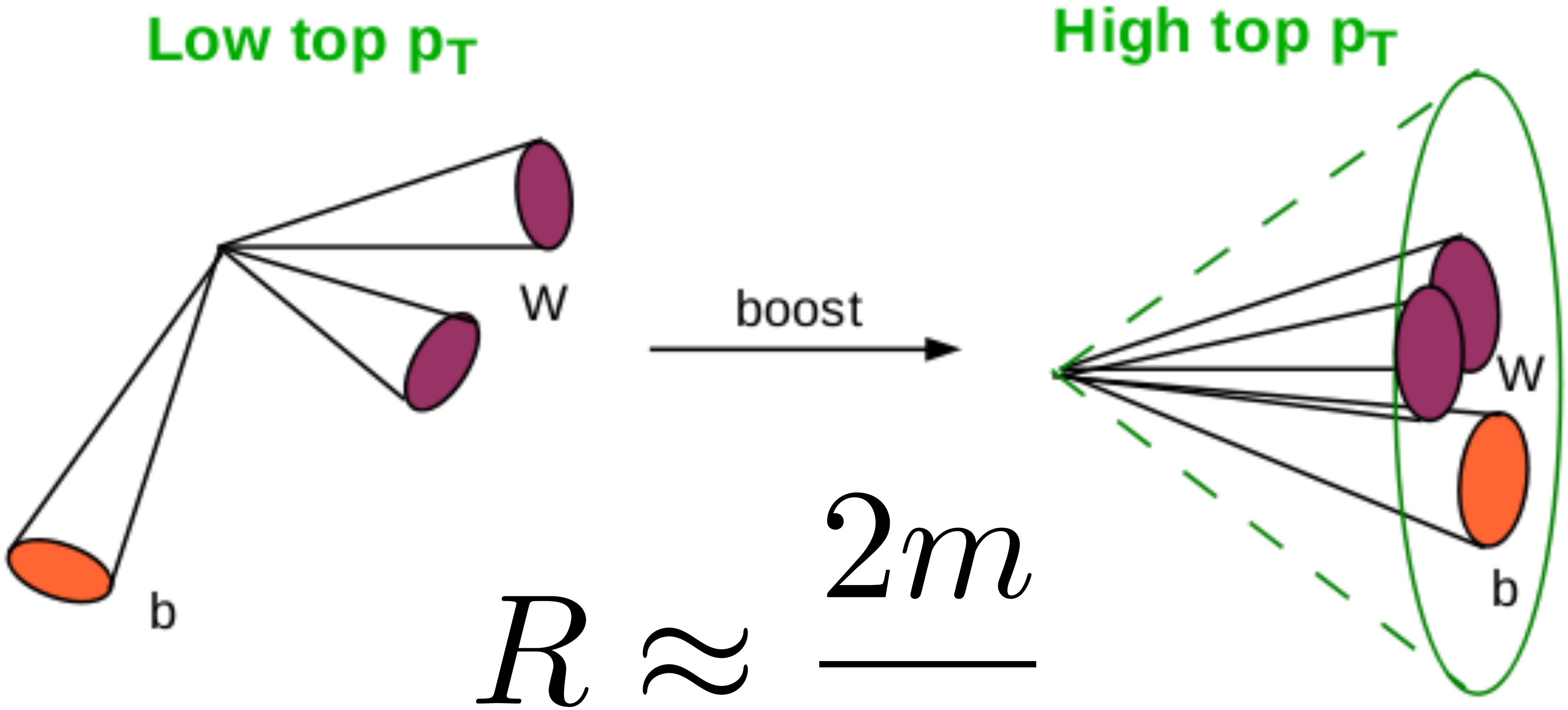
Jet Substructure

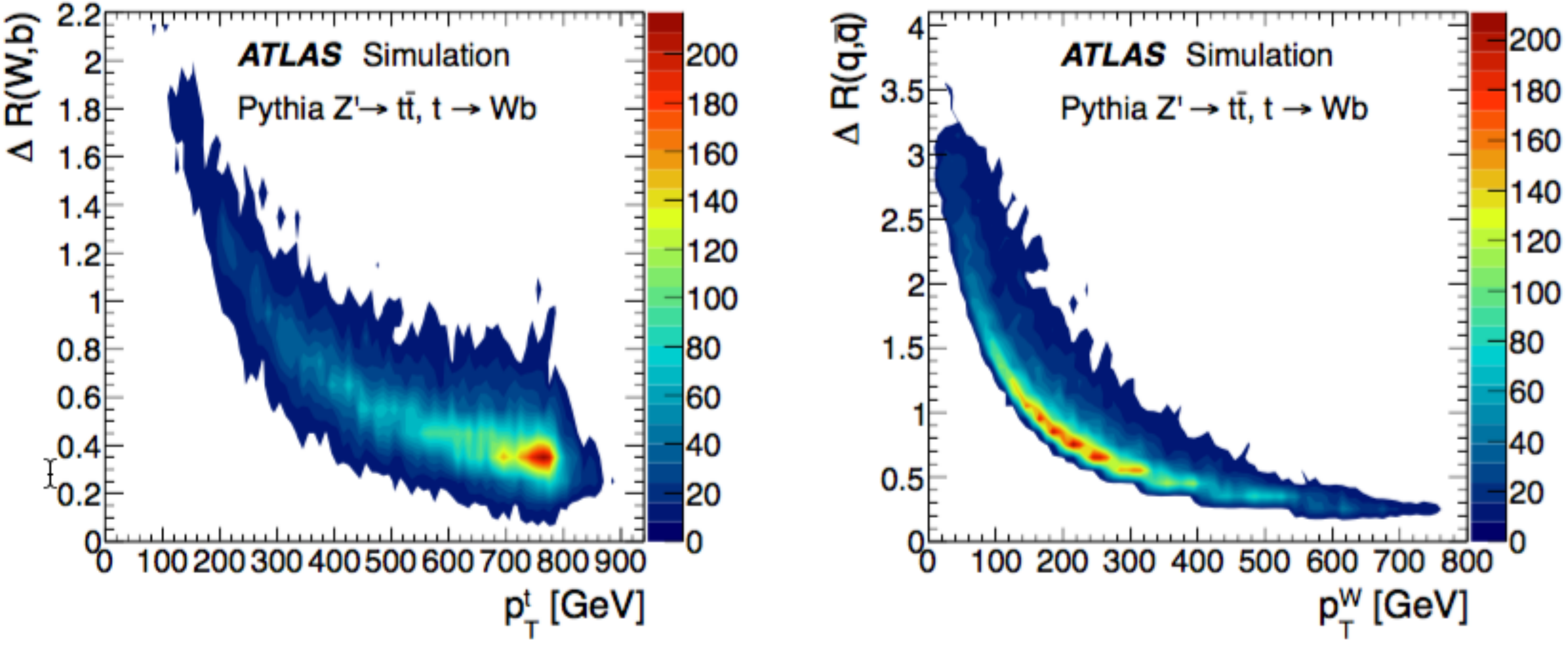
4/27/2016
Physics 290E
Brian Amadio

As the center-of-mass energy of the LHC increases, it becomes harder to resolve the decay products of increasingly boosted SM particles.



$$R \approx \frac{2m}{p_T}$$

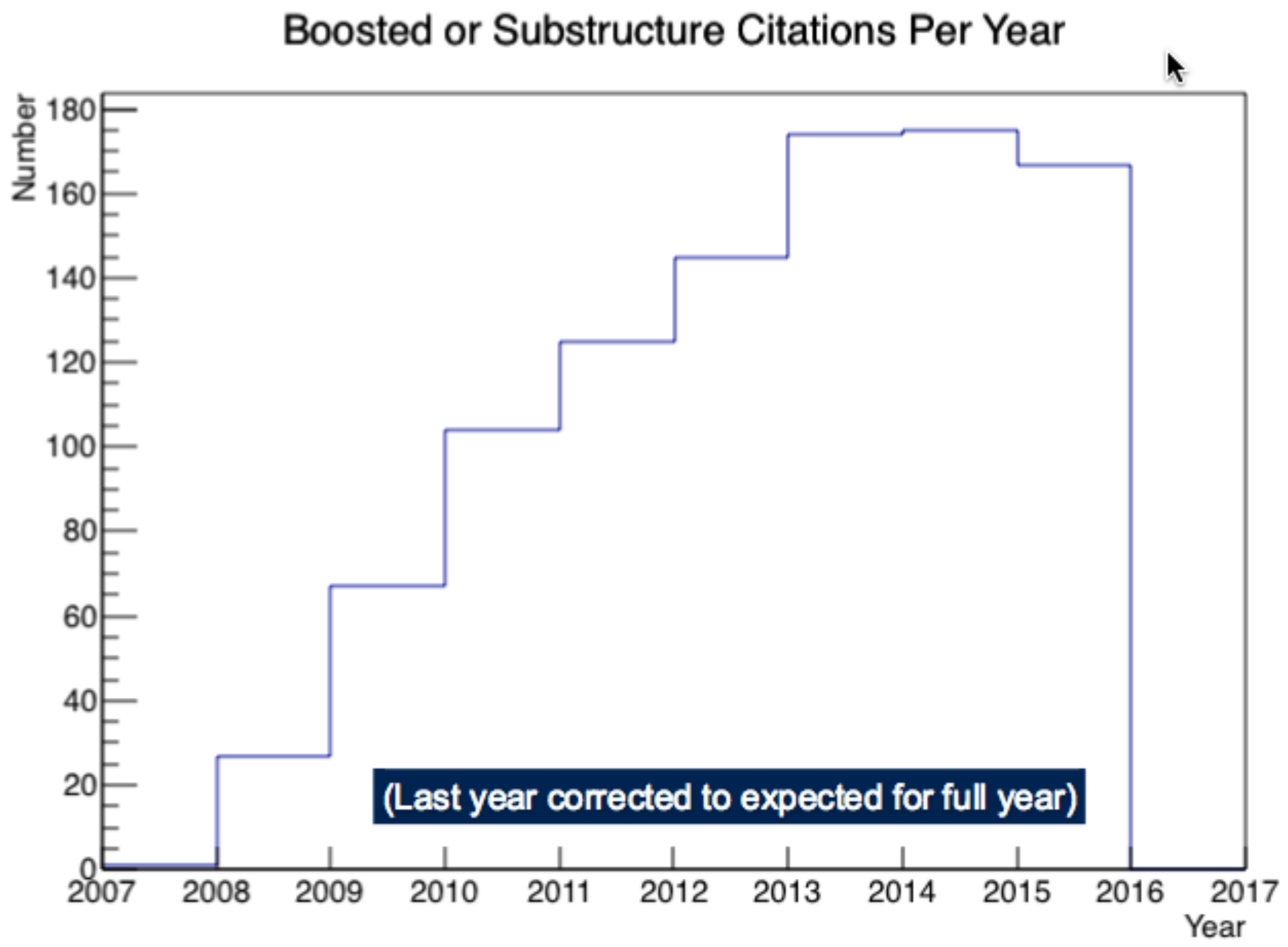
For 1.6 TeV Z'→ tt, very few of the top quarks can be fully resolved with R=0.4 jets



(a) $t \to Wb$

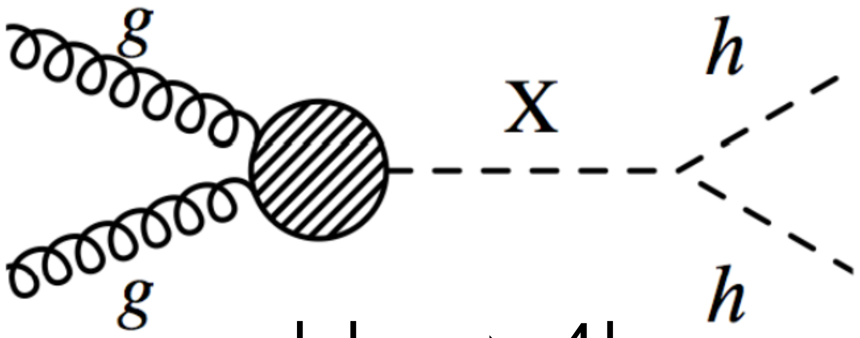
(b) $W \to q\bar{q}$

Collider energies and heavy resonance limits are going up → The future of jets lies in jet substructure!

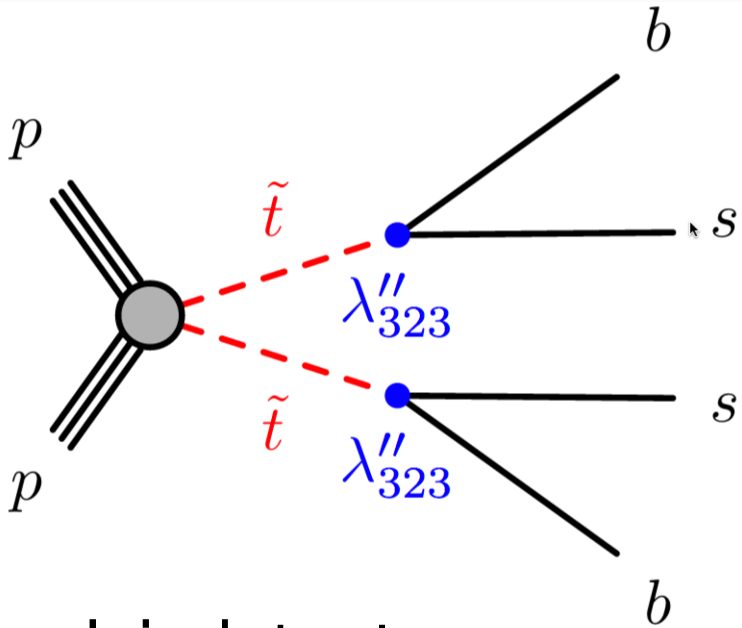


From BOOST 2015 Summary Talk

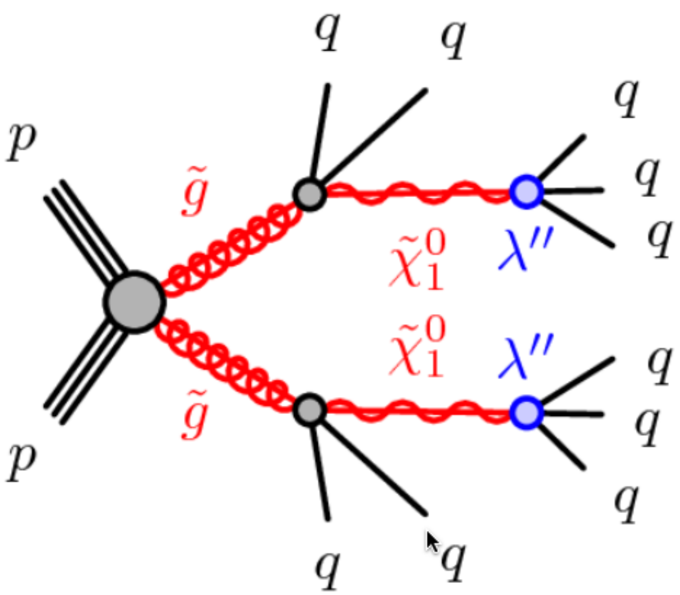
Many BSM searches benefit from jet substructure methods.



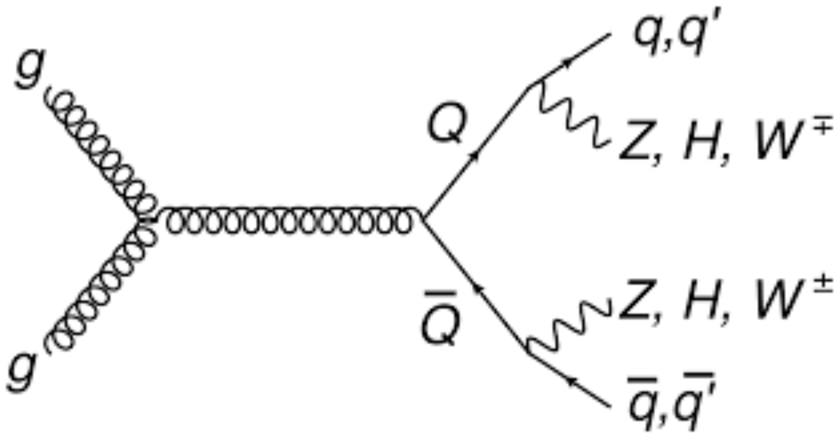
hh → 4b



Light stops

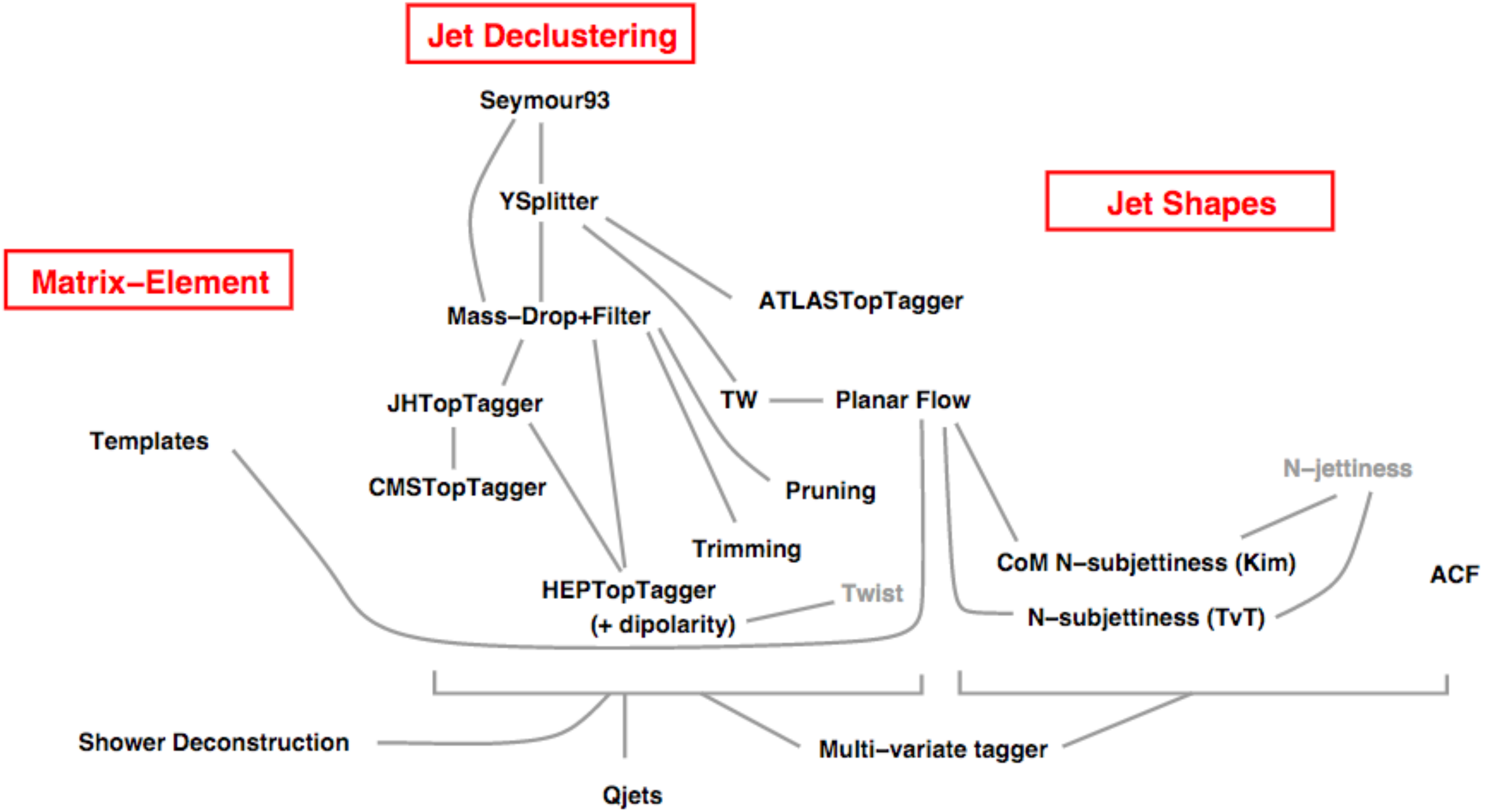


Multijet RPV
Accidental substructure



Vector-like quarks

Jet substructure is a wide, active, and rapidly expanding field.



The most straightforward jet substructure variable is mass.

$$m^2 = \left(\sum E_i\right)^2 - \left(\sum \vec{p}_i\right)^2$$

Boosted resonance: $m_{\text{jet}} \approx m_{\text{resonance}}$

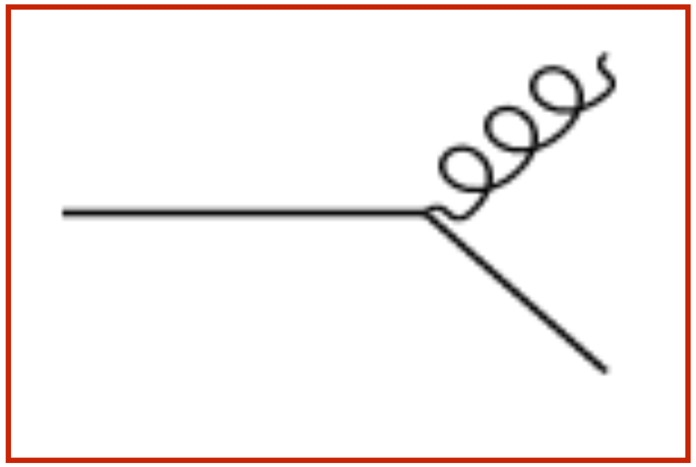
QCD jet: ???

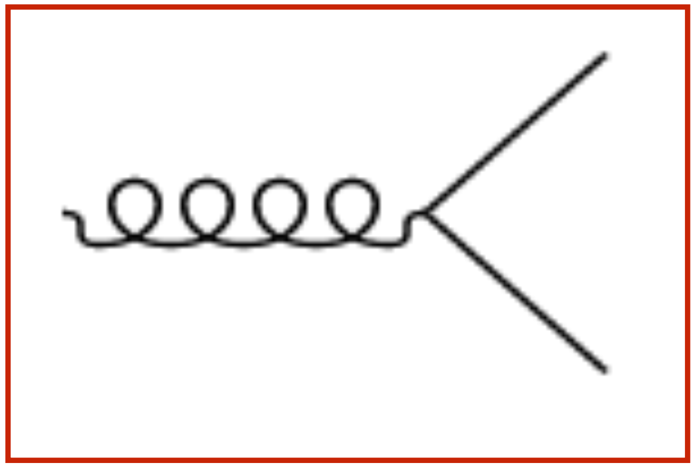
QCD jets are not massless!

$$d\sigma_{n+1} \approx d\sigma_n dz \frac{dt}{\boxed{t}} \frac{\alpha_s}{2\pi} \mathcal{P}(z)$$

t = virtuality = m²

$$\langle m^2 \rangle \approx p_{T,P}^2 \int_0^{R^2} \frac{d\theta^2}{\theta^2} \int dz \;\; z\,(1-z)\,\theta^2 \frac{\alpha_s}{2\pi} \mathcal{P}(z)$$





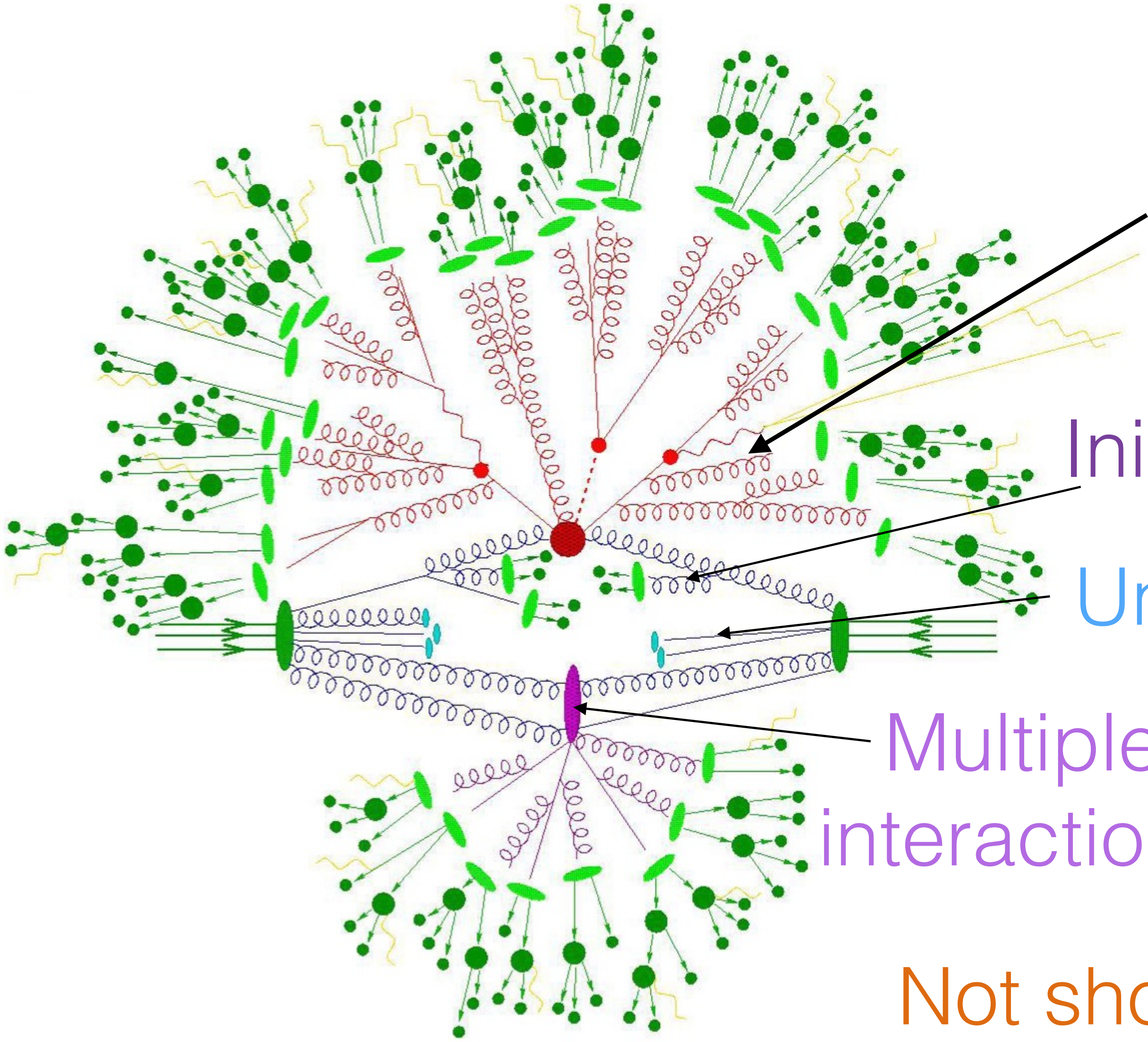
$$\langle m^2 \rangle \approx \frac{\alpha_s}{\pi} \frac{3}{8} C_F p_T^2 R^2$$

quark jet

$$\langle m^2 \rangle \approx \frac{\alpha_s}{\pi} \frac{1}{20} C_A p_T^2 R^2$$

gluon jet

For finite-radius jets, the mass is sensitive to unassociated radiation in the event.



Final State Radiation

Initial State Radiation

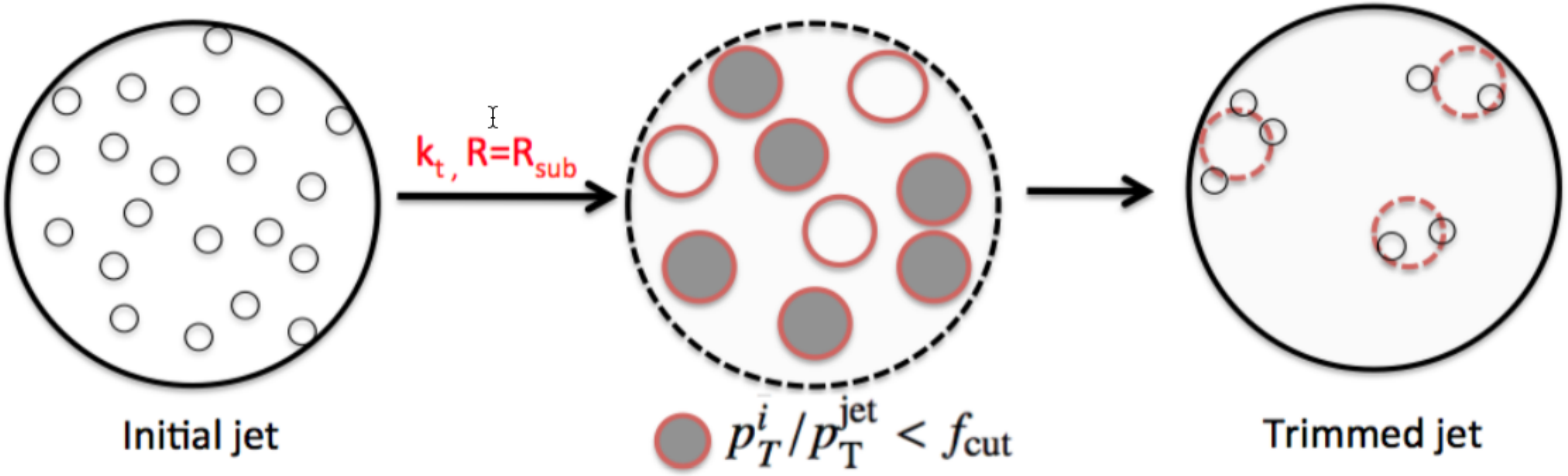
Underlying event

Multiple interactions

$$m^2 = \left(\sum E_i\right)^2 - \left(\sum \vec{p}_i\right)^2$$

Not shown: pileup

Trimming is one of several jet grooming methods used to reduce this sensitivity.



Unassociated radiation should be ≈evenly distributed in the event, while radiation from parent partons is clustered near the parent. Use k_T for reclustering to keep more FSR.

Mass-drop/filtering takes advantage of the fact that large changes in mass are exponentially suppressed in QCD showers.

Sudakov form factor $$\Delta\left(t\right) = exp\left[-\int_{t_0}^{t}\frac{dt'}{t'}dz\frac{\alpha_s}{2\pi}\mathcal{P}\left(z\right)\right]$$

$$\Delta\left(t\right) \propto \left(\frac{t_0}{t}\right)^p$$

But when a heavy particle decays, large changes in mass are guaranteed!

Mass-drop/filtering attempts to optimize the reclustering scale for a specific decay topology, originally h→bb.

Algorithm: unwind C/A jets and require:

1. large mass drop $max(m_1, m_2) < \mu m_J$

   and

2. symmetric splitting $\frac{min\left(p_{T1}^2, p_{T2}^2\right)}{m_J^2} \Delta R_{1,2}^2 > y_{cut}$

If either requirement fails, throw out softer subjet and unwind harder subjet with same method. Continue until a splitting passes both requirements to obtain R_{filt} = min(0.3,ΔR/2) Then recluster N subjets at R_{filt}.

Pruning, inspired by filtering, explicitly vetoes soft, wide angle radiation.

Algorithm: unwind C/A jets and require:

1. hard splitting $$\frac{min\left(p_{Ti}, p_{Tj}\right)}{p_{TJ}} > z_{cut}$$

 or

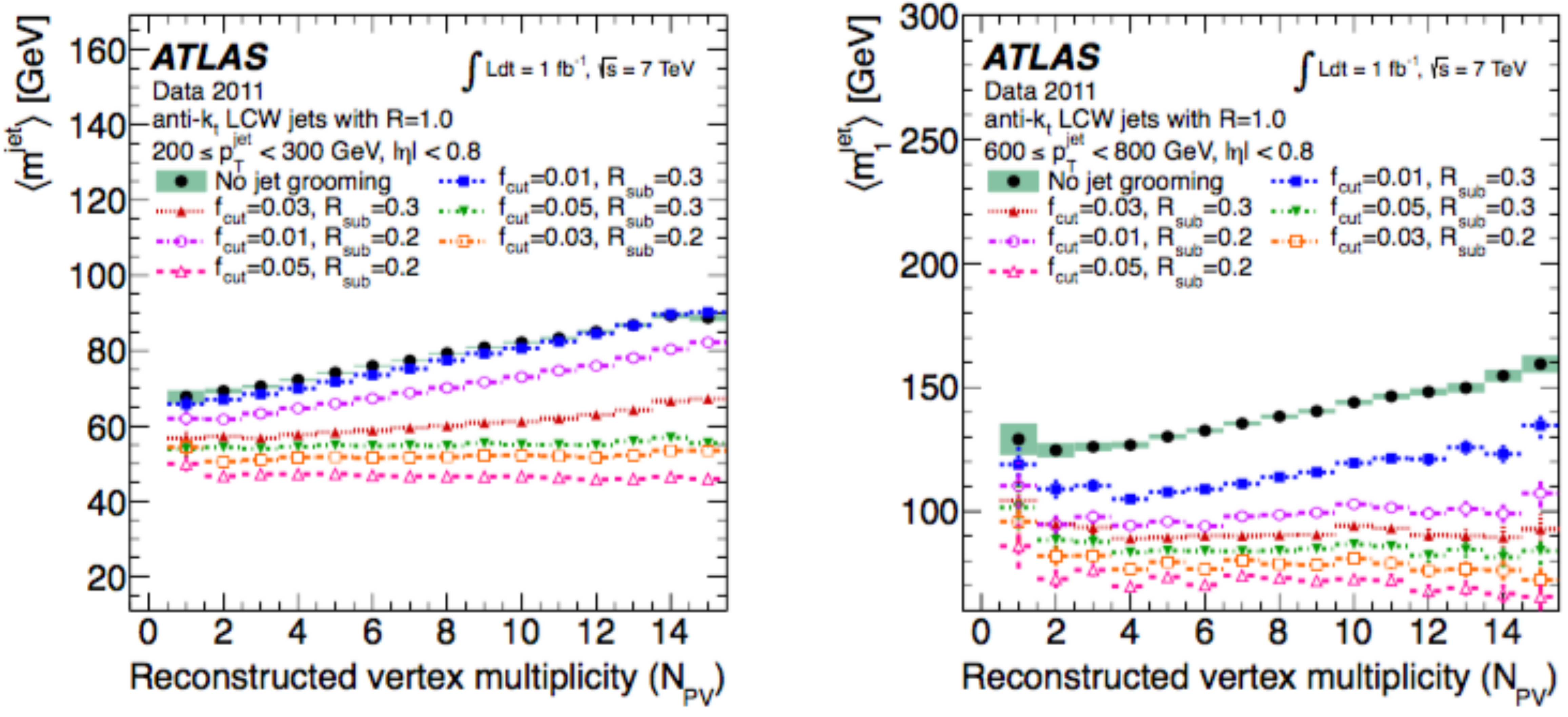
2. small angle $$\Delta R_{ij} < D_{cut}$$

At each step, if both checks fail, drop the softer constituent and unwind the harder. Stop when either 1 or 2 are satisfied.

All grooming methods come with parameters that have to be tuned.

grooming method	trimming	mass-drop	pruning
parameters	R, R_{sub}, f_{cut}	R, μ, y_{cut}	R, z_{cut}, D

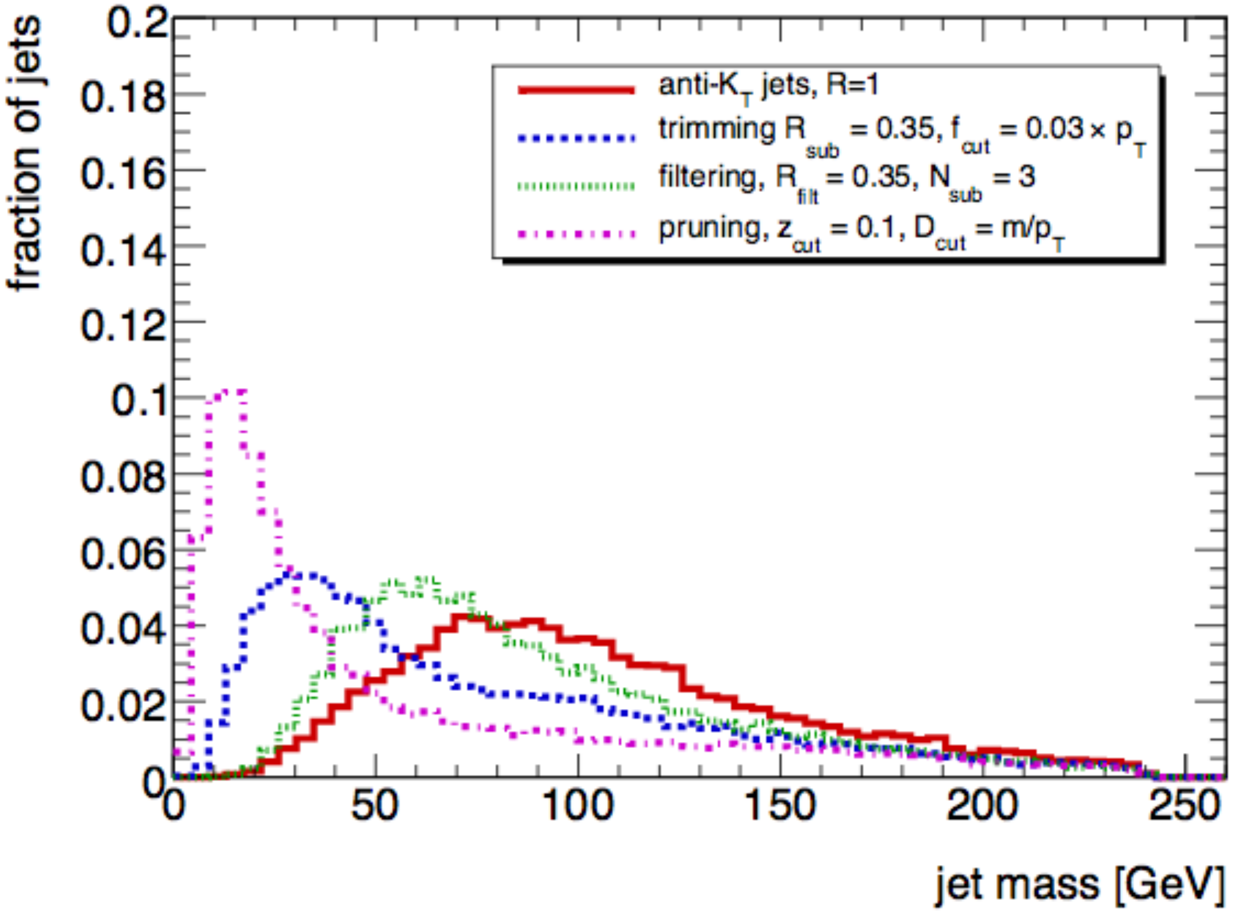
Grooming can significantly reduce pileup sensitivity.



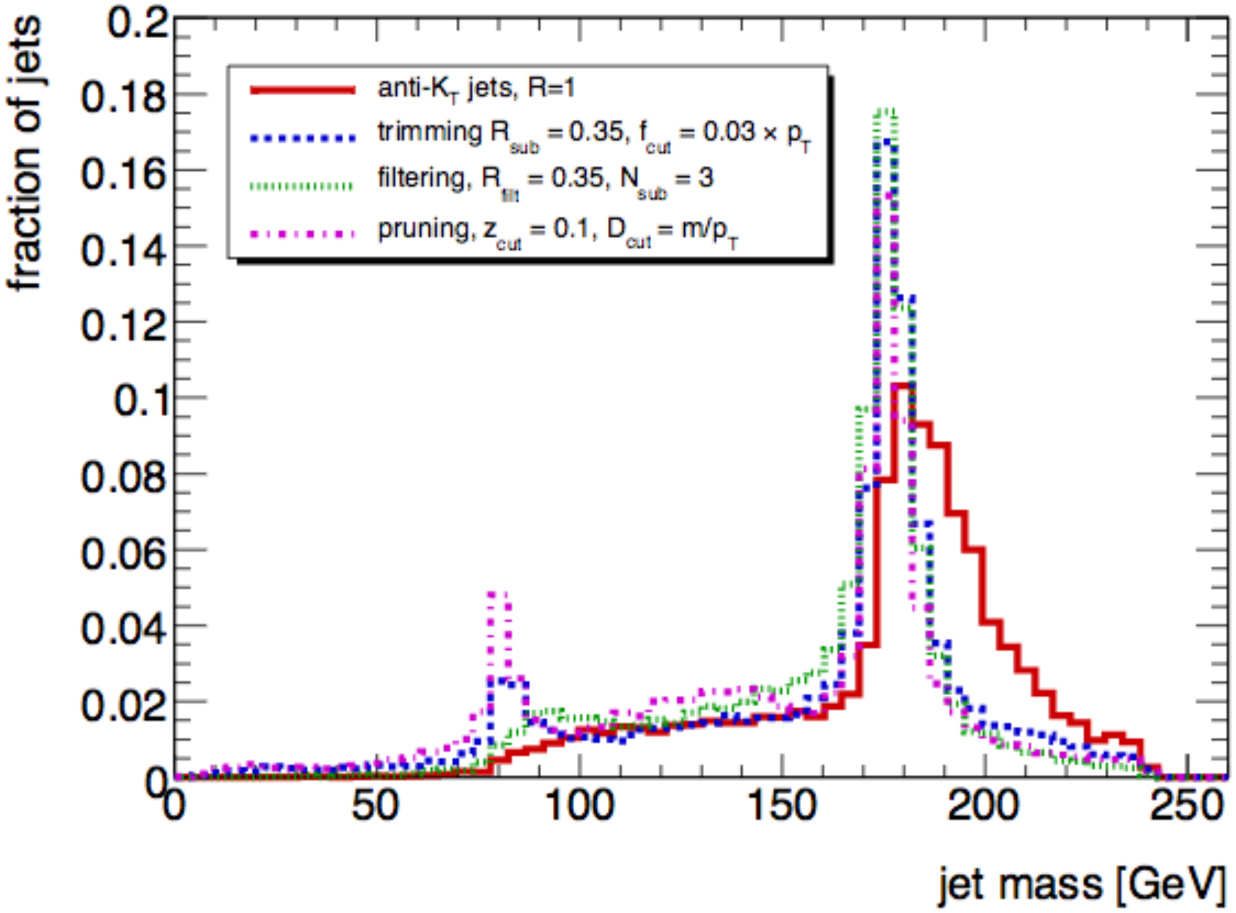
(a) Trimmed anti-k_t: $200\ \mathrm{GeV} \leq p_T^{\mathrm{jet}} < 300\ \mathrm{GeV}$

(b) Trimmed anti-k_t: $600\ \mathrm{GeV} \leq p_T^{\mathrm{jet}} < 800\ \mathrm{GeV}$

Grooming improves the jet mass resolution for signal and reduces the QCD background.



(a) dijets, 500–600 GeV



(b) $t\bar{t}$, 500–600 GeV

Pruning most aggressive at small mass, as D scales like m/p_T, whereas R_{sub} is fixed. Filtering least aggressive, as it keeps a fixed number of subjets - in this case 3.

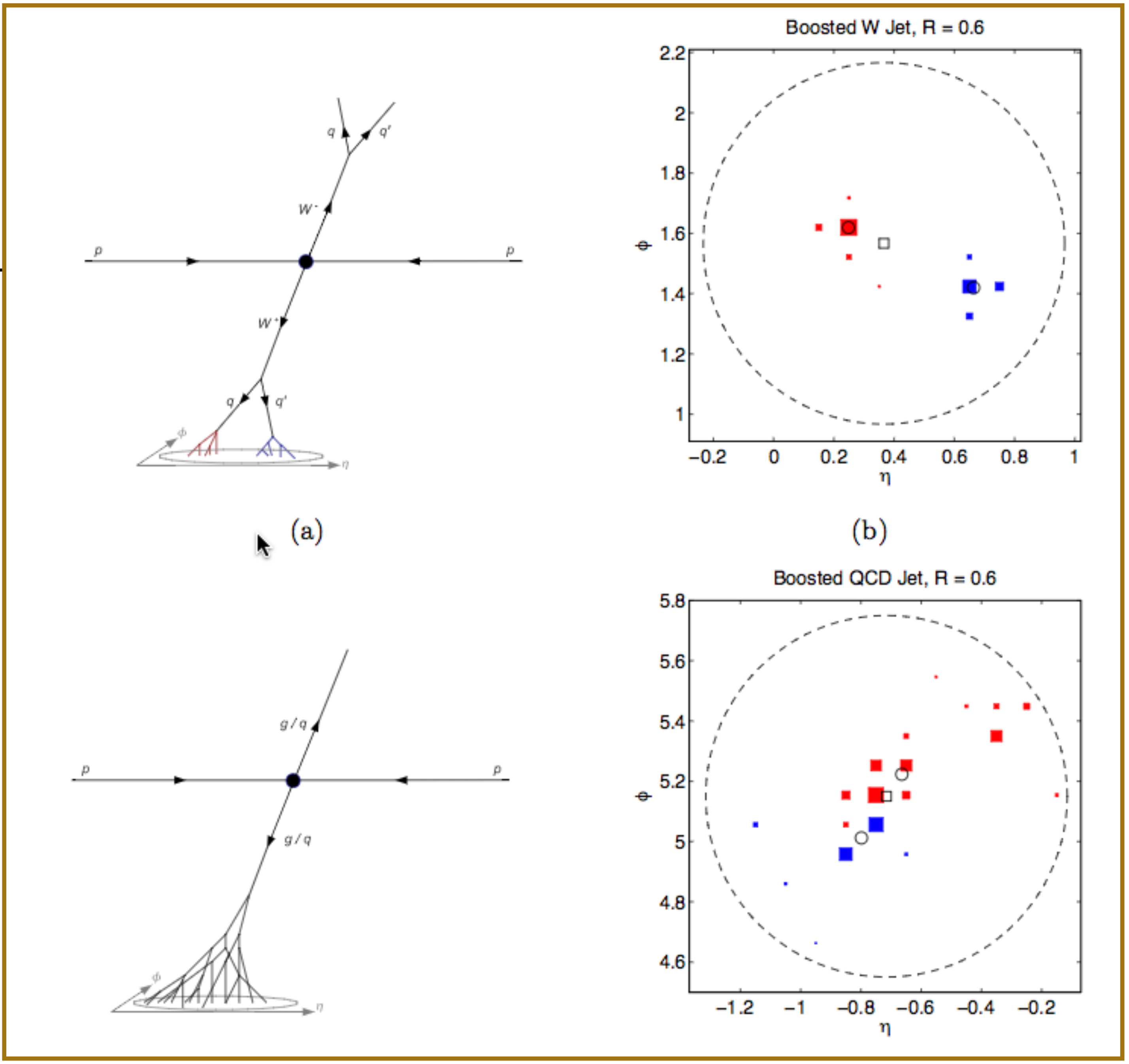
N-subjettiness quantifies the degree to which a jet is comprised of N subjets.

Given N axes $\hat{n}_k$,

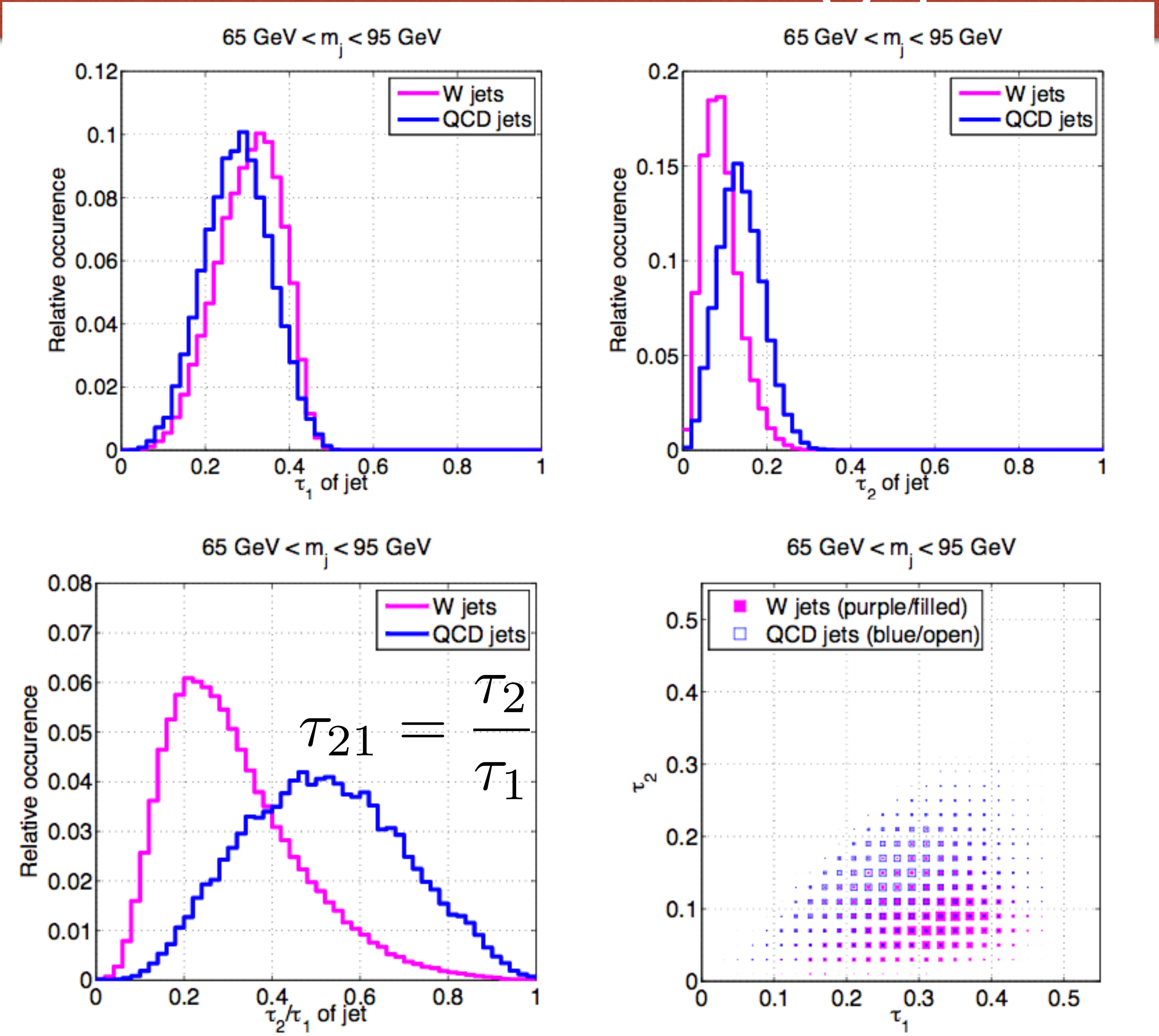
$$\tau_N = \frac{\sum_{i \in J} p_{T,i} min\left(\Delta R_{ik}\right)}{\sum_{i \in J} p_{T,i} R_0}$$

τ_N small: jet is well described by N or fewer subjets

τ_N large: more subjets will likely describe jet better



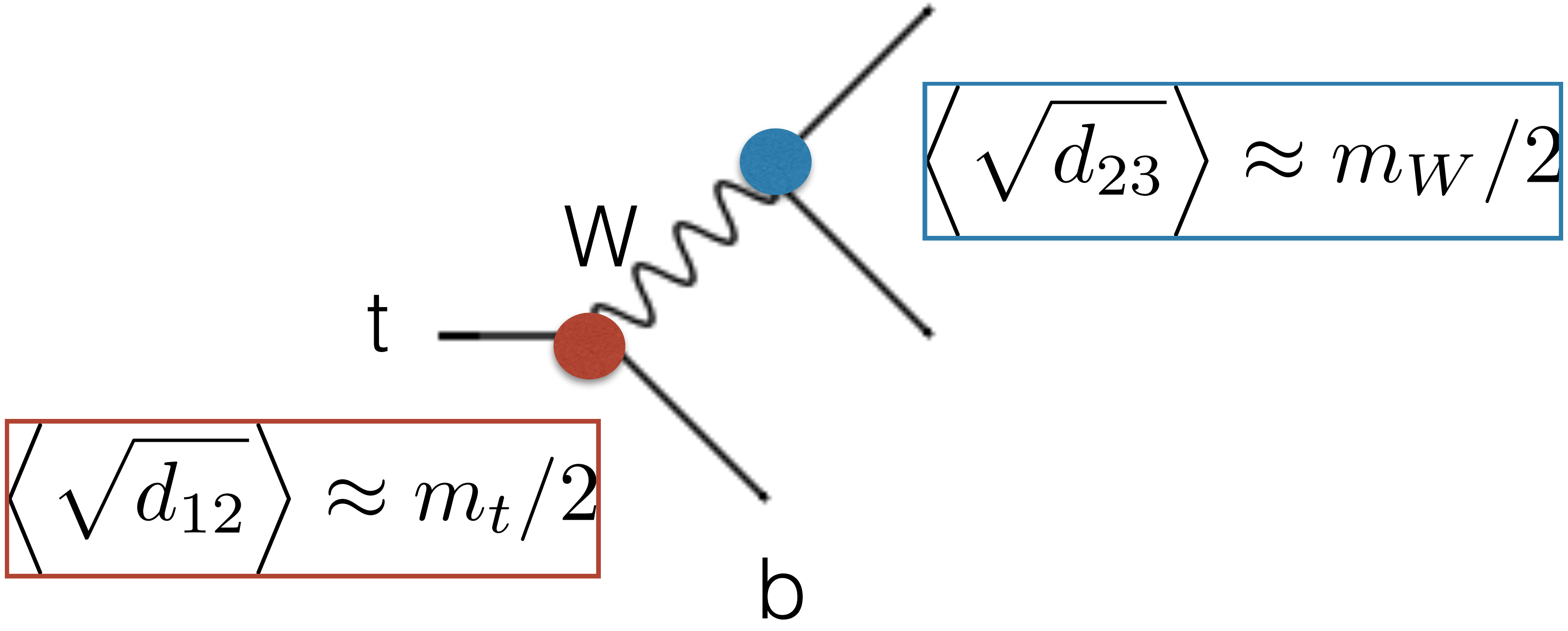
Ratios of N-subjettiness provide more discriminating power.



$$\tau_{21} = \frac{\tau_2}{\tau_1}$$

The kT splitting scale can also be used to separate QCD jets from boosted resonance decay jets.

Recluster jets using k_T algorithm $\sqrt{d_{ij}} = min\left(p_{Ti}, p_{Tj}\right)\Delta R$

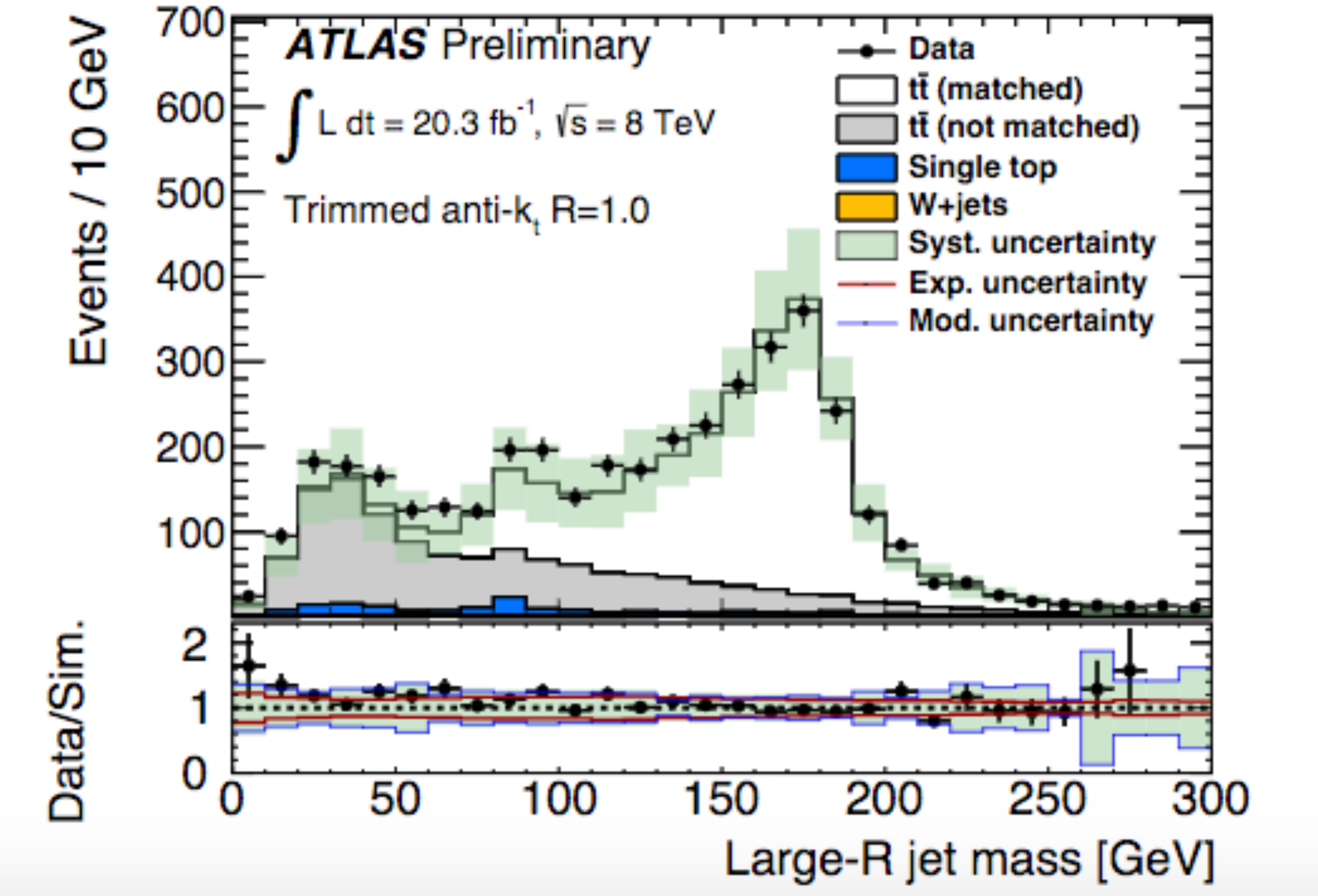


$$\left\langle\sqrt{d_{23}}\right\rangle \approx m_W/2$$

$$\left\langle\sqrt{d_{12}}\right\rangle \approx m_t/2$$

For QCD jets: $d_{34} \ll d_{23} \ll d_{12}$

To study top-tagging performance, we use a semileptonic ttbar selection.



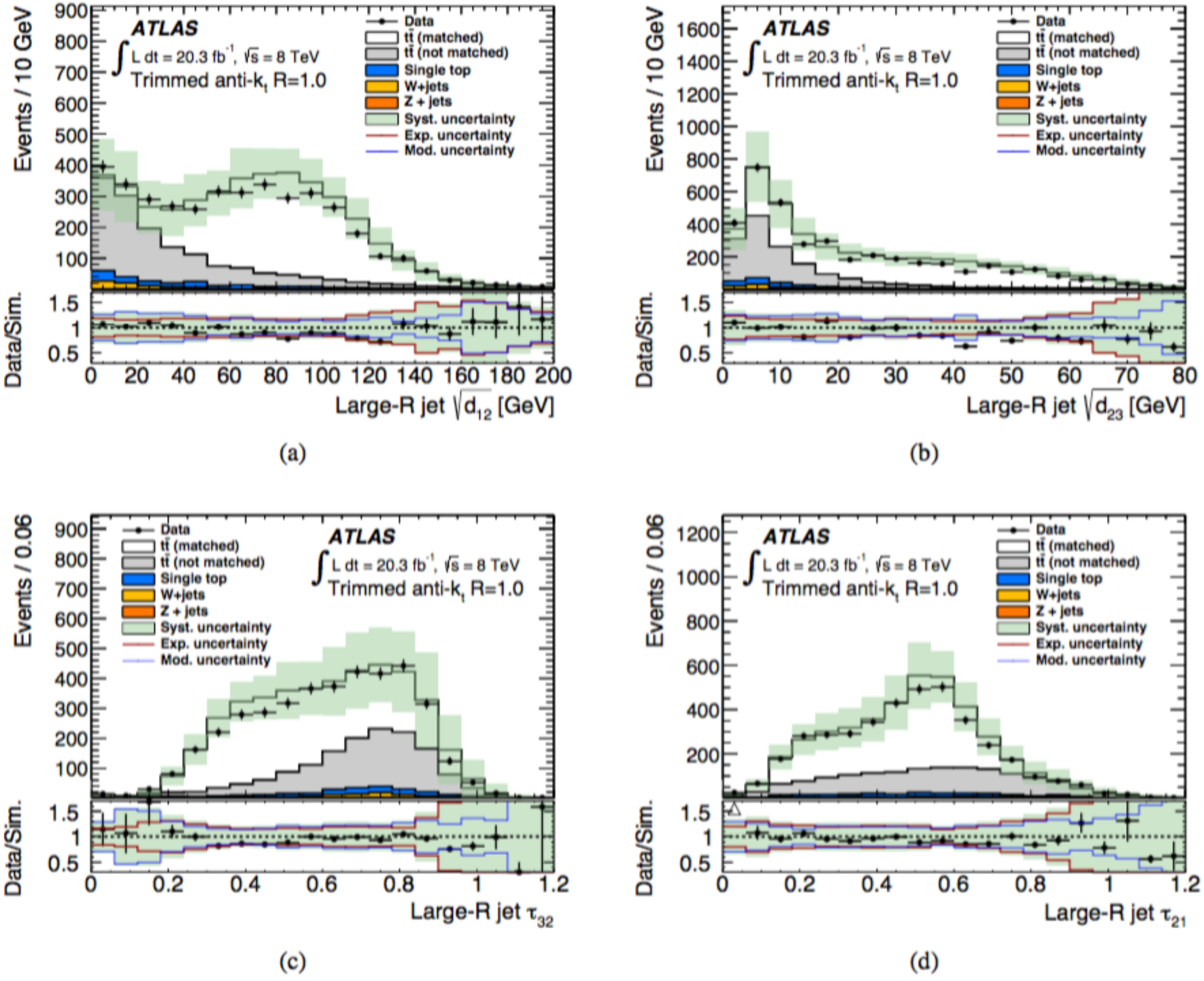
Selection

1 lepton, 2 b-tagged jets.

Truth-matching

large-R jet within ΔR = 0.75 of hadronically-decaying top at decay vertex.

Substructure-based top-tagging can combine mass, n-subjettiness, and k_T splitting scales.



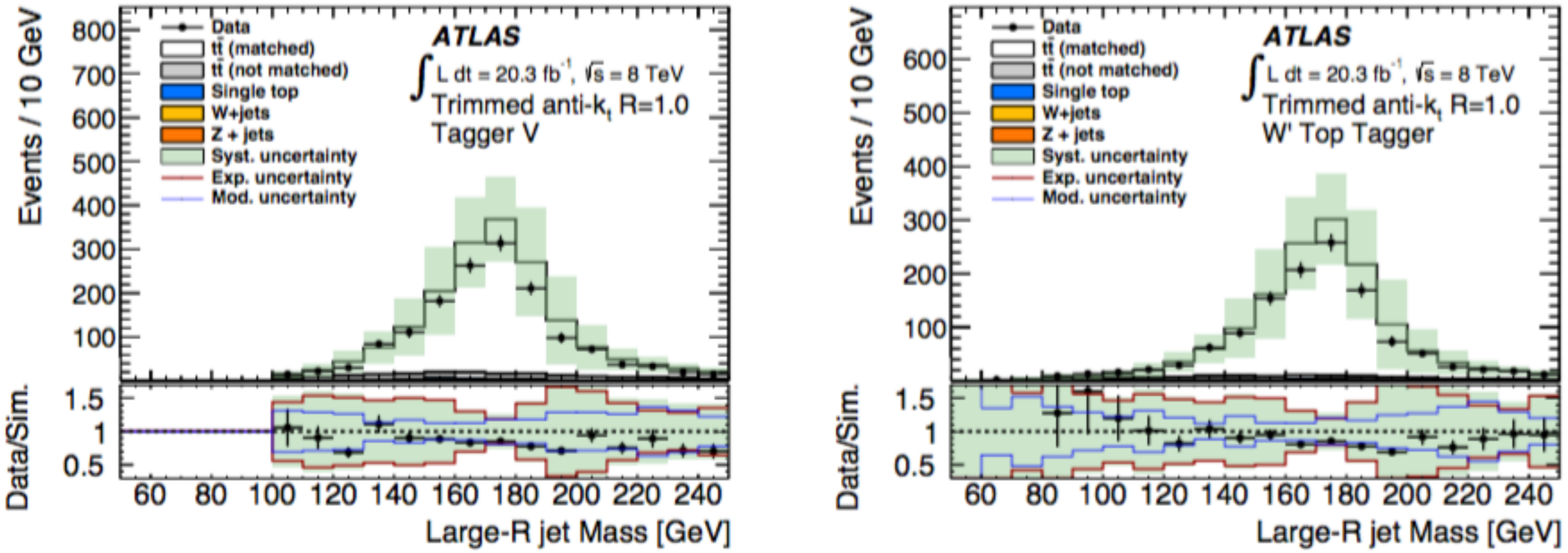
(a)

(b)

(c)

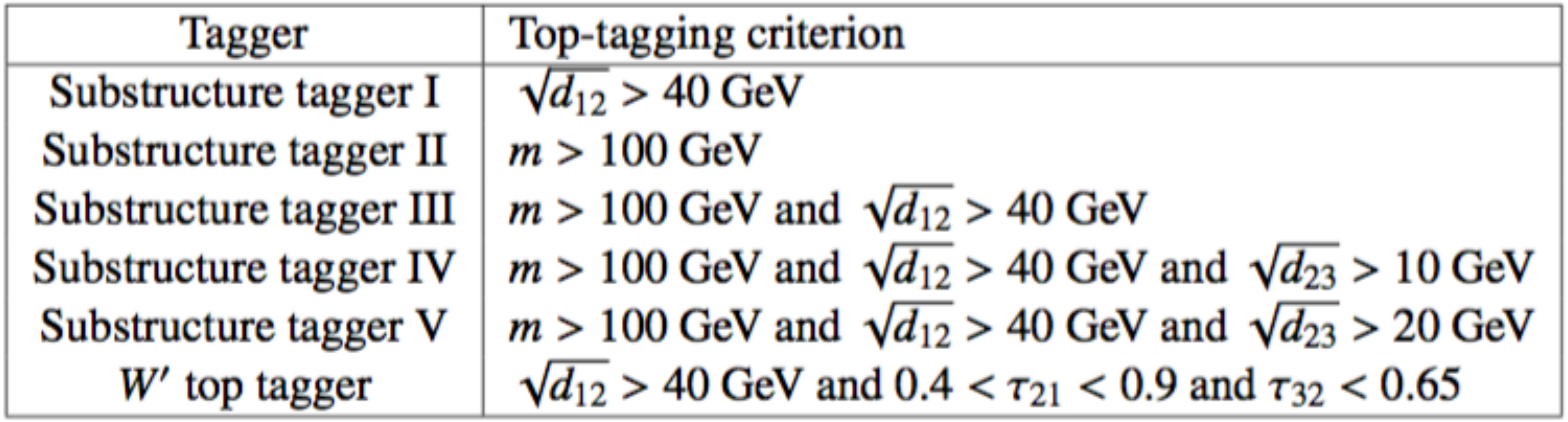
(d)

The signal-to-background is improved as compared to mass cuts alone.



(e) Tagger V

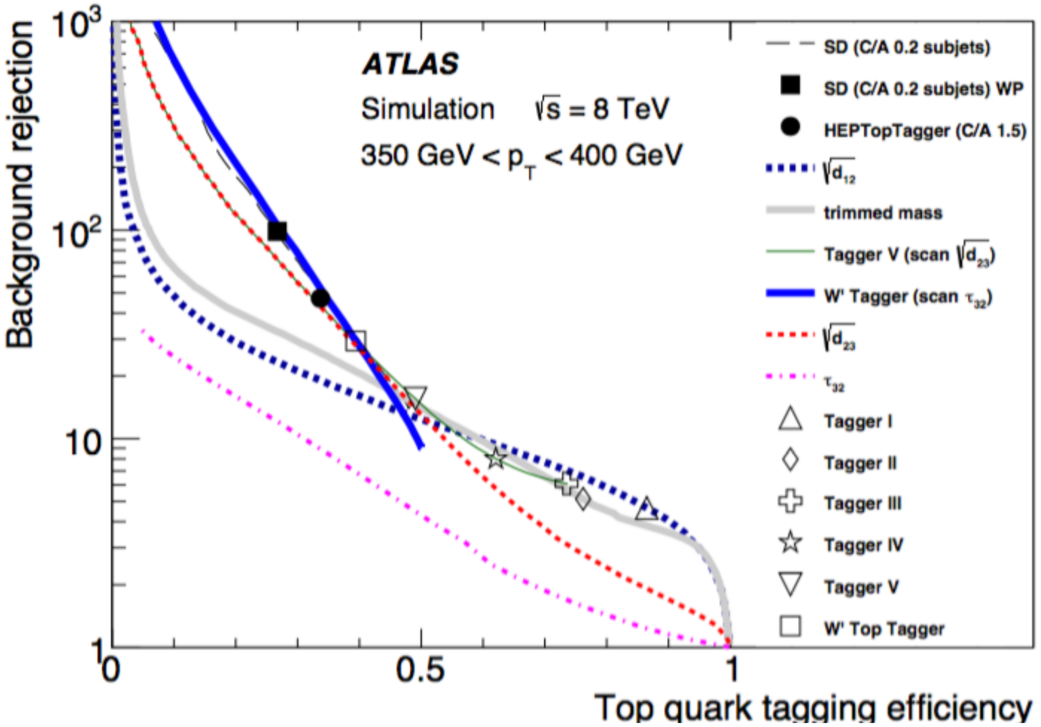
(f) W' top tagger

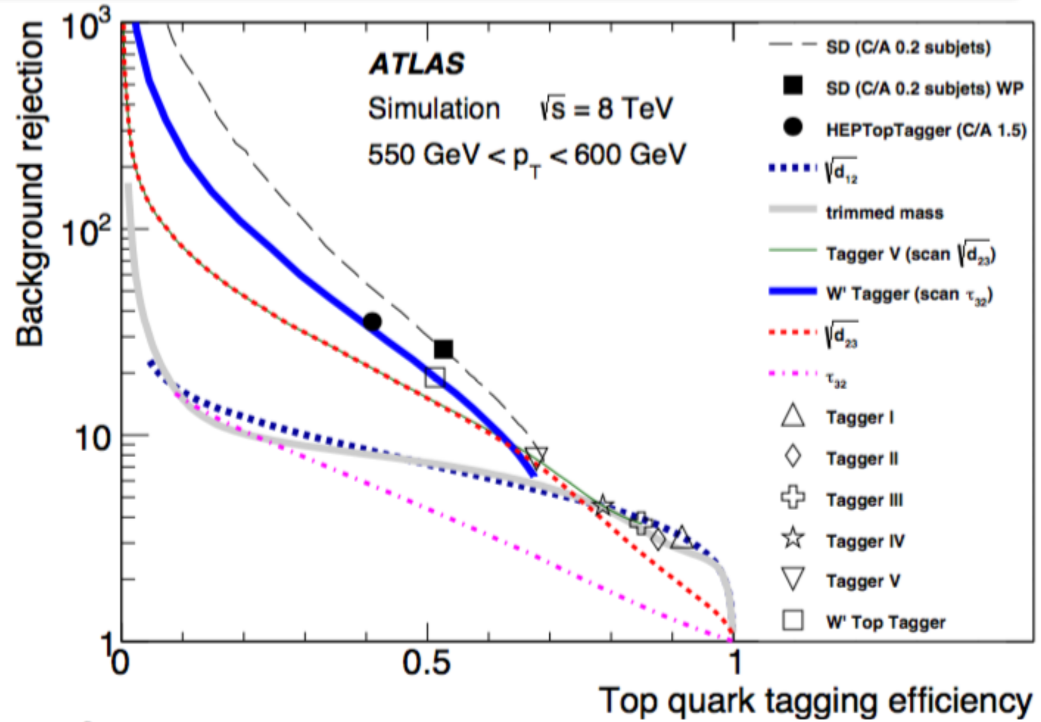


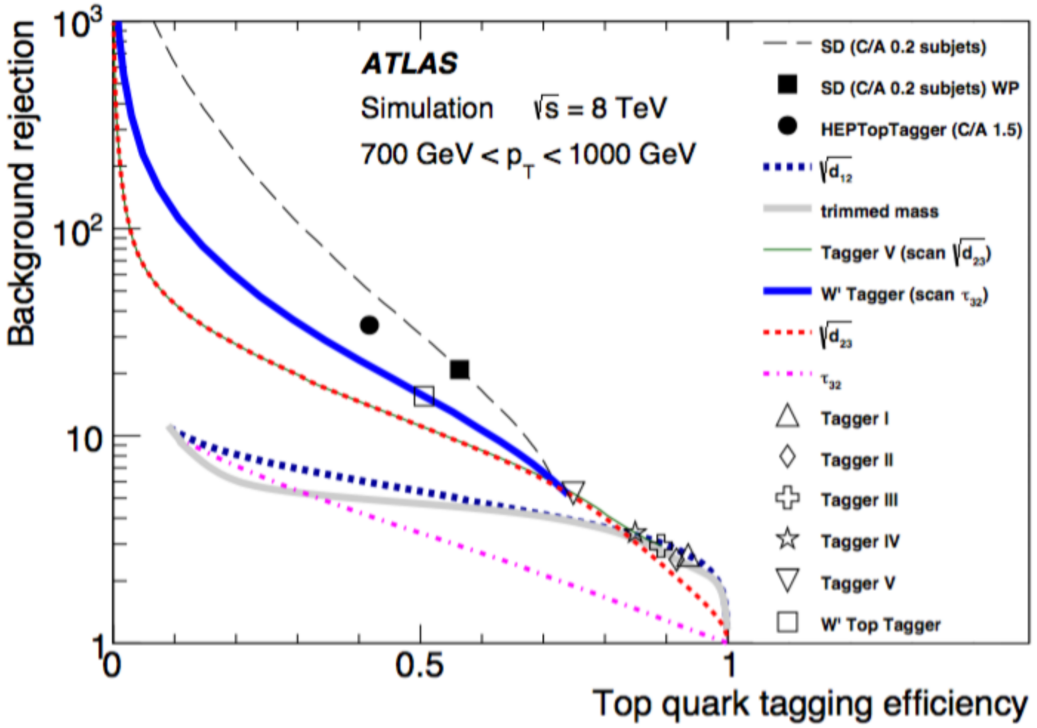
Tagger	Top-tagging criterion
Substructure tagger I	$\sqrt{d_{12}} > 40$ GeV
Substructure tagger II	$m > 100$ GeV
Substructure tagger III	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV
Substructure tagger IV	$m > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV and $\sqrt{d_{23}} > 10$ GeV
Substructure tagger V	$m > 100$ 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 $\tau_{32} < 0.65$

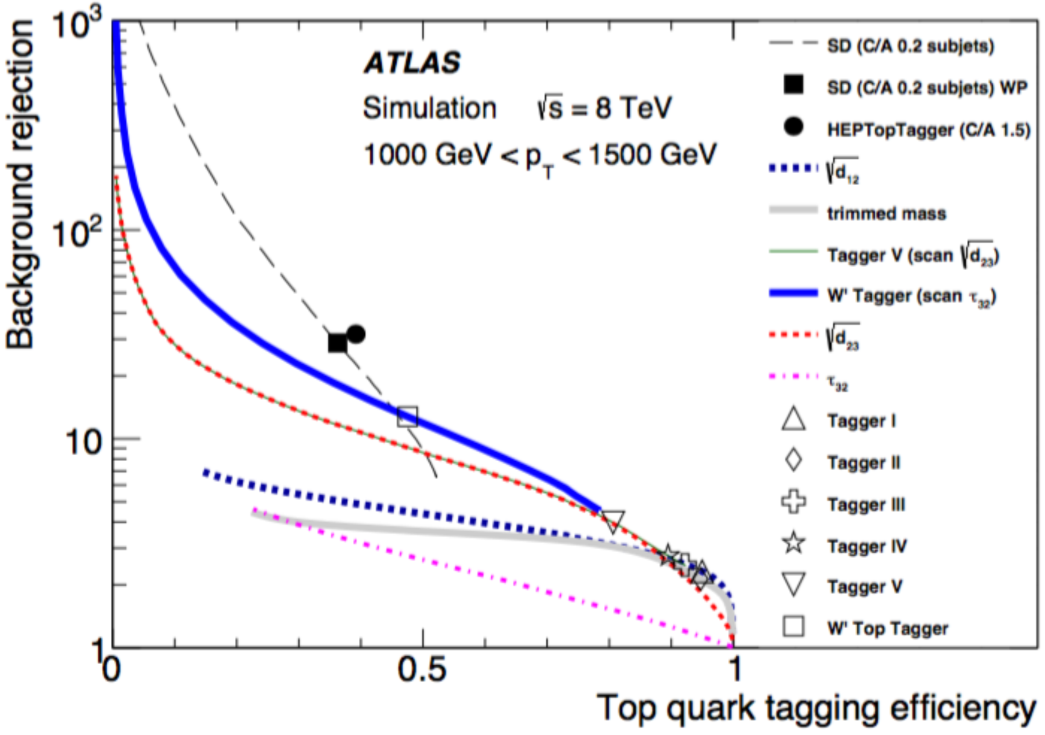
Table 2: Top taggers based on substructure variables of trimmed anti-k_t $R = 1.0$ jets.

Performance is studied in MC, using Z’→ ttbar for signal and multijets for background.









Additional Material

HEPTopTagger

Use Cambridge-Aachen, $R = 1.5$ jets

Algorithm

1. unwind the jet and require $\max(m_1,m_2) > 0.8\, m_{jet}$ to keep both constituents, otherwise keep higher mass
2. keep unwinding all subjets as long as $m > 30$ GeV. Keep all subjets that haven't been discarded by requirements in step 1

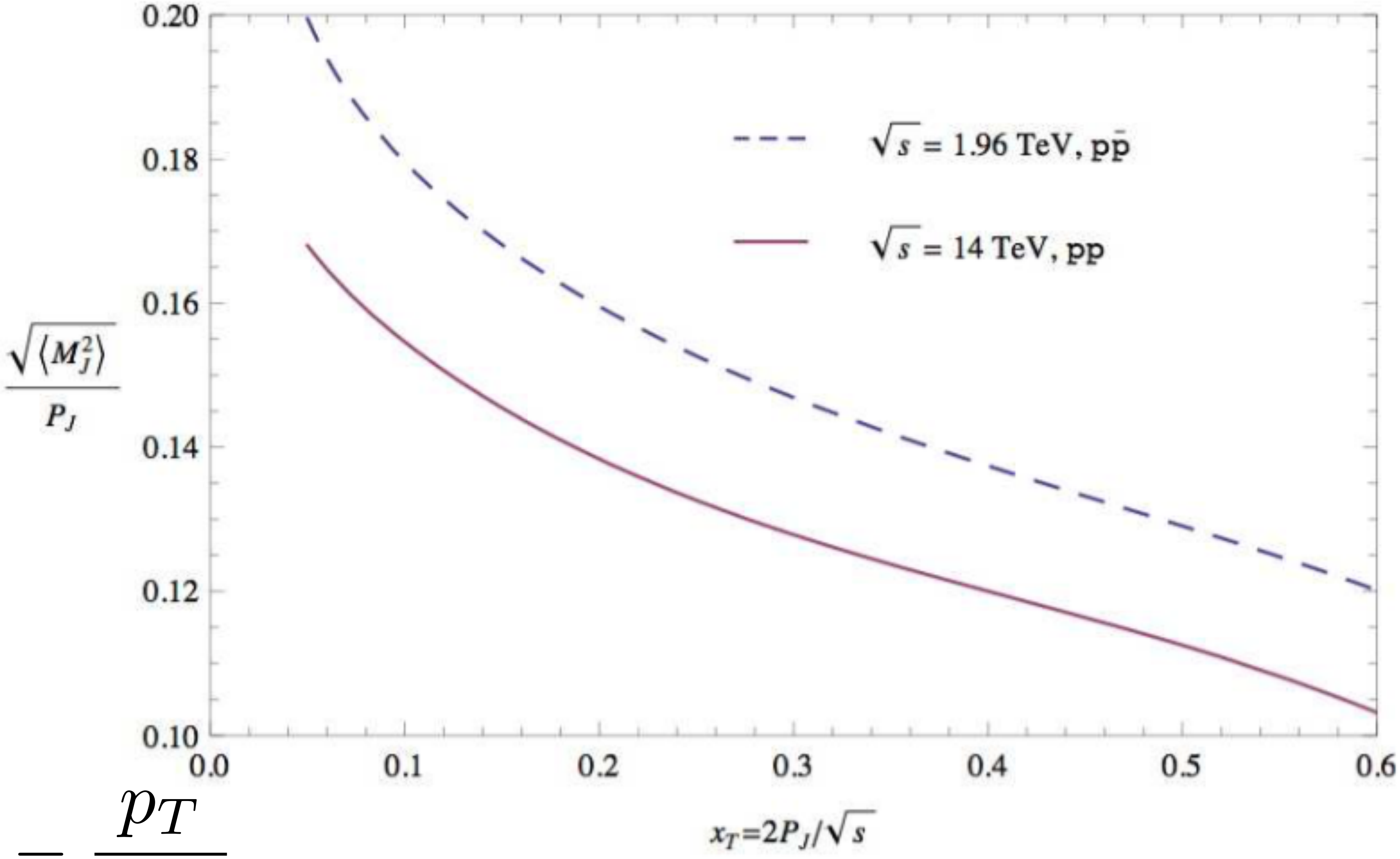
cluster subjets with $R_{filt} = \min(0.3, \Delta R_{jk}/2)$

calculate mass of 5 hardest constituents in each subjet

choose pairings of subjets to obtain total mass closest to m_t

Because of changes in q/g, α_s, and pdfs, QCD jet mass is not quite linear in p_T

NLO calculation for 2 partons in a jet



$$\frac{\sqrt{\langle m^2 \rangle}}{p_T R} \approx 0.25 - \frac{p_T}{3\sqrt{s}}$$

Better than 25%