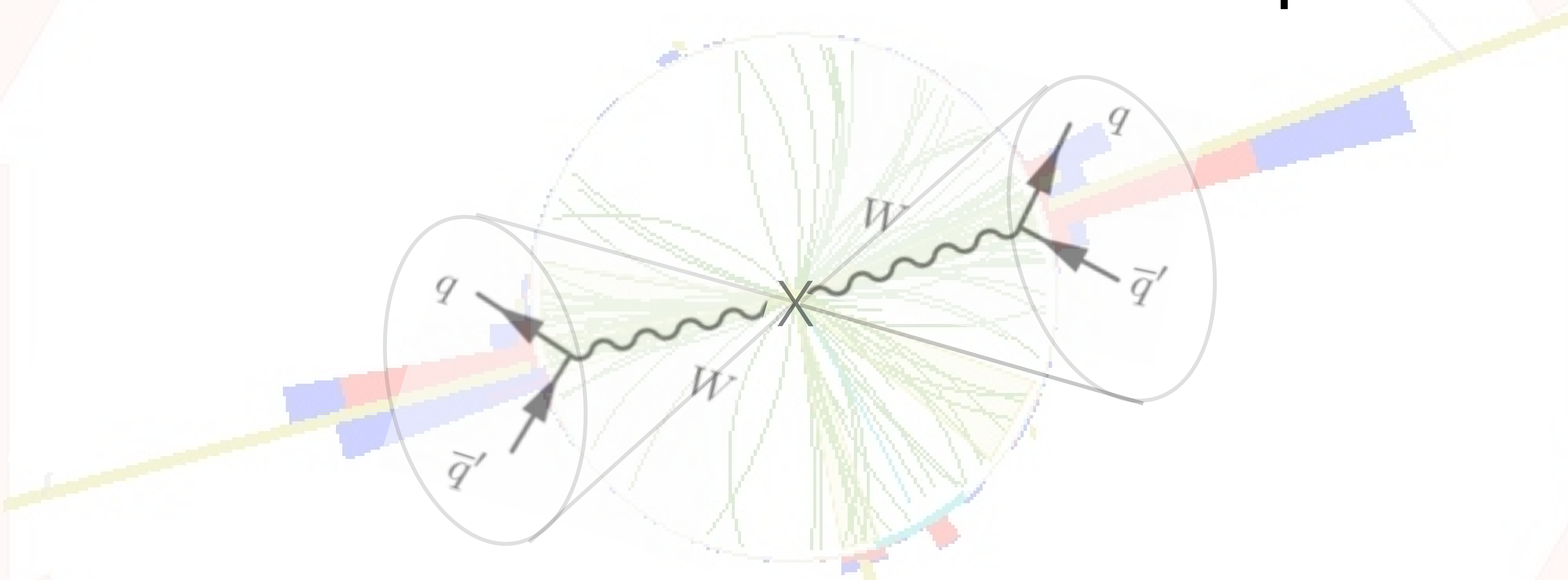
Jet substructure techniques



Thea Klaeboe Aarrestad

UZH CMS seminar, PSI 15.10.15

Physics Institute , University of Zurich

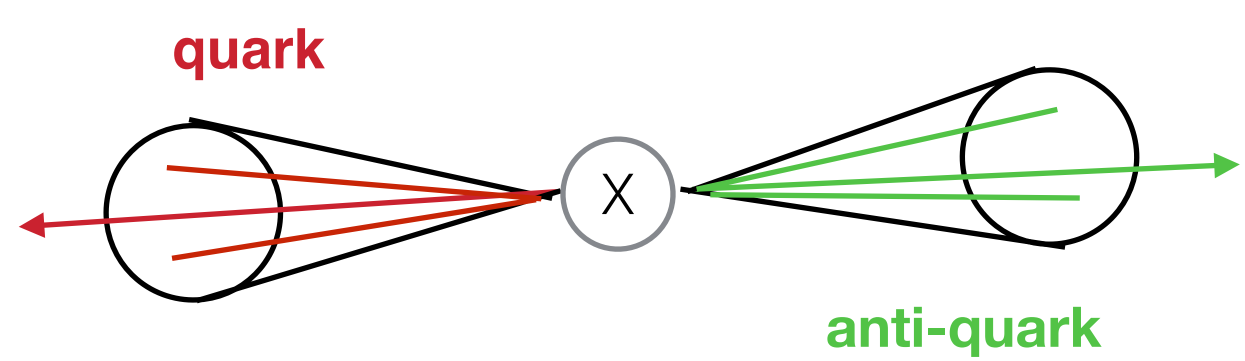
Overview

- Boosted topologies
- Jet substructure techniques
  - *Problems and solutions*
  - *Grooming*
  - *N-subjettiness*
  - *Tagging*
- Jet substructure at CMS
  - *Top tagging*
  - *W/Z tagging*

Boosted topologies

- Normal analyses: two quarks from $X \rightarrow q\bar{q}$ reconstructed as two jets

Low boost



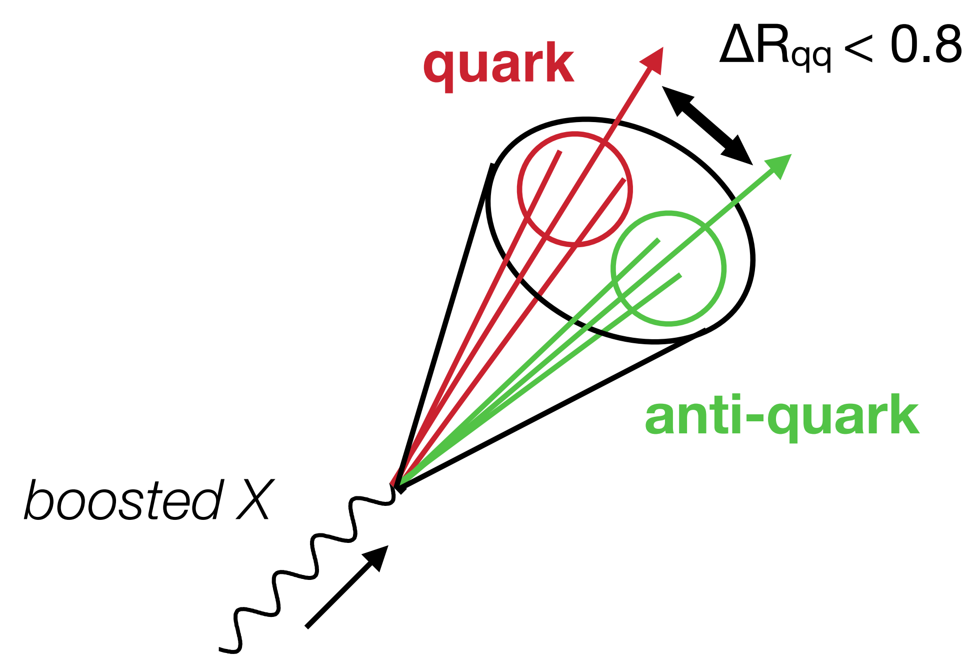
Boosted topologies

Universität Zürich[UZH]

- Normal analyses: two quarks from $X \rightarrow q\bar{q}$ reconstructed as two jets
- High-p_T regime: Object X is boosted, decay is collimated, $q\bar{q}$ both in same jet!

$$\Delta R_{qq}^{\min} \approx \Delta\theta_{qq}^{\min} \approx 2\frac{M_V}{p_{T,V}}$$

High boost



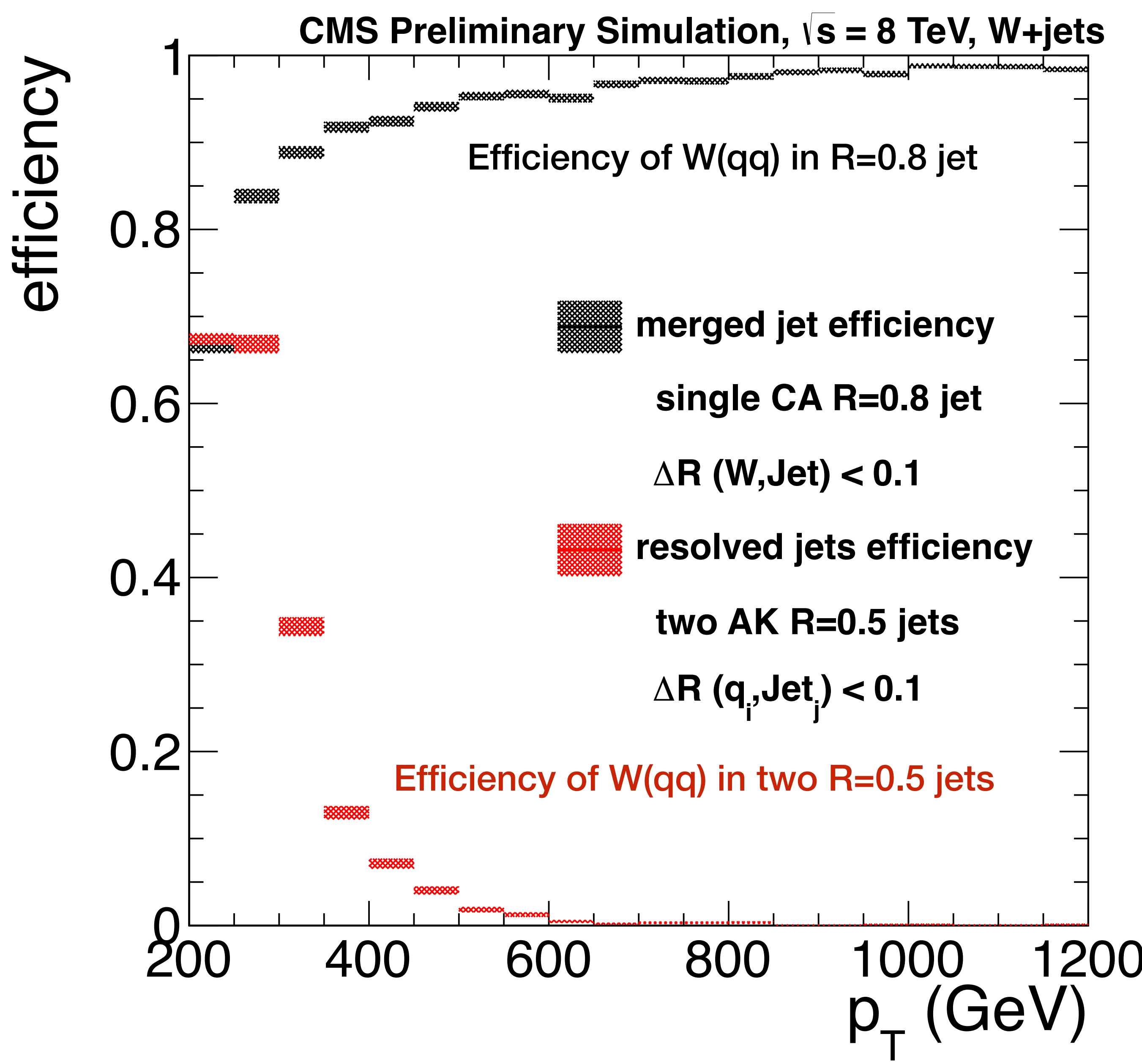
Boosted topologies

Universität Zürich[UZH]

- Normal analyses: two quarks from $X \rightarrow q\bar{q}$ reconstructed as two jets
- High-pt regime: EW object X is boosted, decay is collimated, $q\bar{q}$ both in same jet!

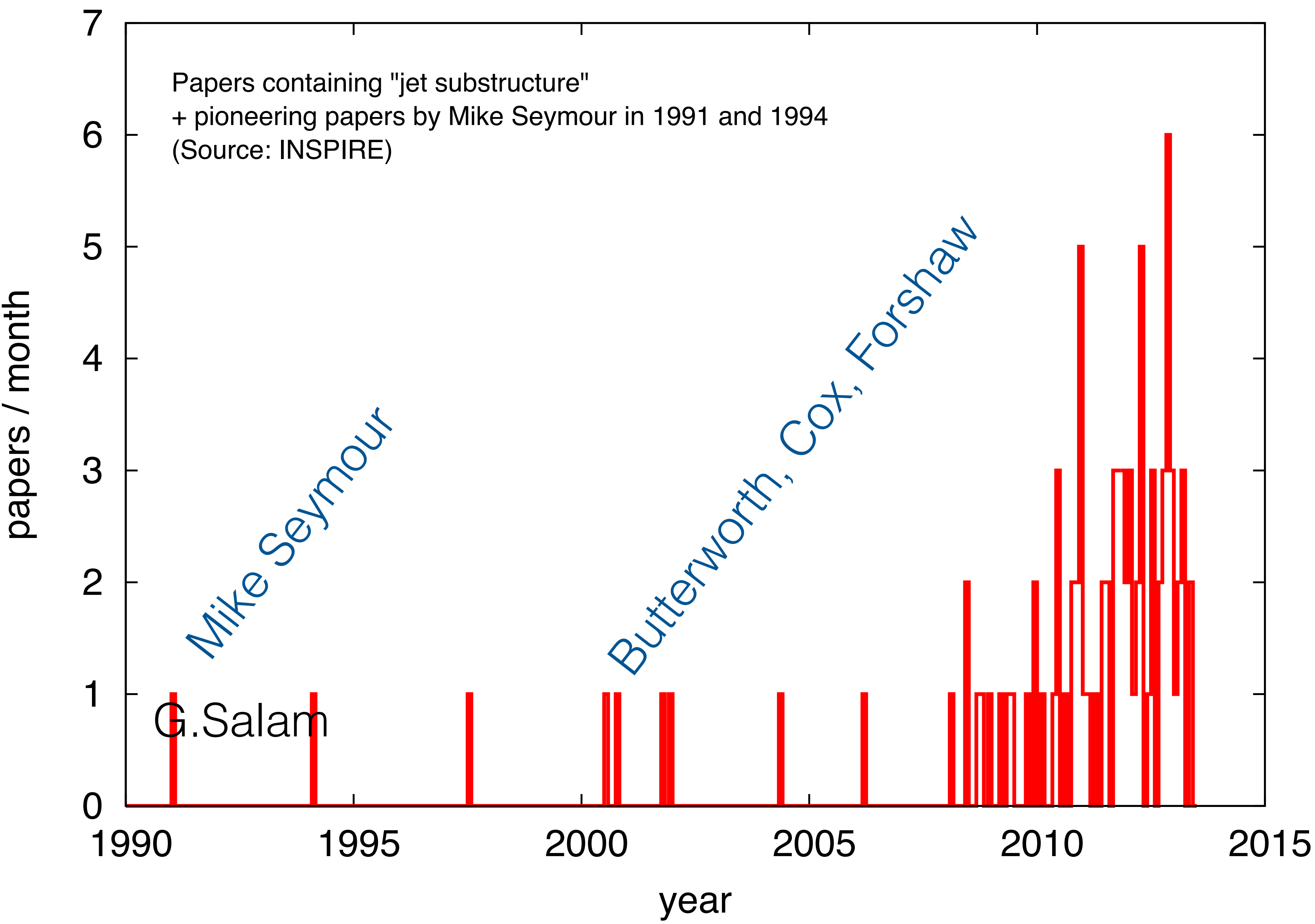
$$\Delta R_{qq}^{\min} \approx \Delta\theta_{qq}^{\min} \approx 2\frac{M_V}{p_{T,V}}$$

- Example: For W/Z emerging from decay of > 1 TeV resonance:
  - *Above W/Z p_T > 200 GeV, decay products merged into single massive R = 0.8 jet!*
- Efficiency to tag energetic W in one large jet way more feasible than using the resolved q jets at high p_T



Jet substructure techniques

Universität Zürich[UZH]



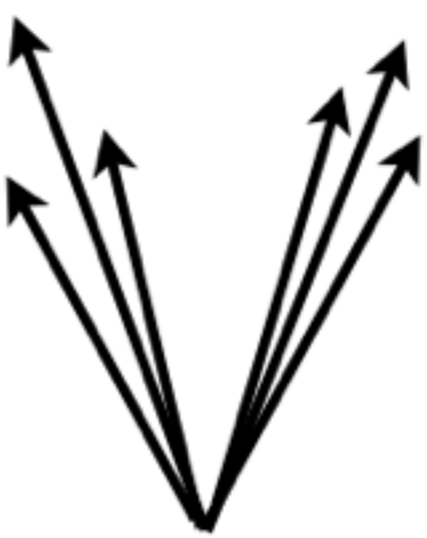
- More than 1000 papers on jet substructure since 2008
- Very active and growing research field!
- Today I will be focusing on techniques used in W/Z/H and top tagging

Jet substructure techniques

Universität Zürich[UZH]

- Jet substructure techniques attempt to step back from the final jet clustering and identify smaller structures within wide jet
- All jet substructure techniques take advantage of properties of signal and background jets related to the jet mass and substructure

SIGNAL PROPERTIES

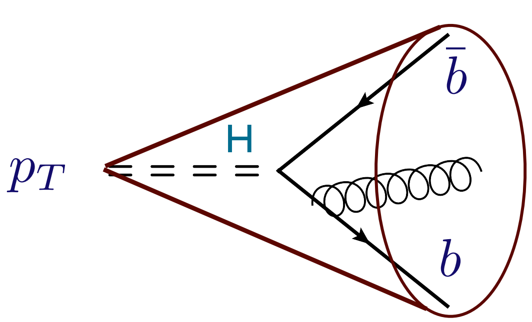


Signal jet

- *Massive 4-vector sum of daughters*
- *1-2 wide-angle splittings*
- *Symmetric splittings*

Jet mass

Jet substructure



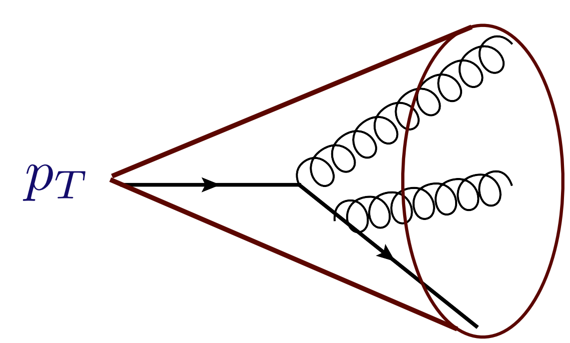
symmetric splitting

$p_T^b \sim p_T^{\bar{b}}$

BACKGROUND PROPERTIES

Quark/gluon jets

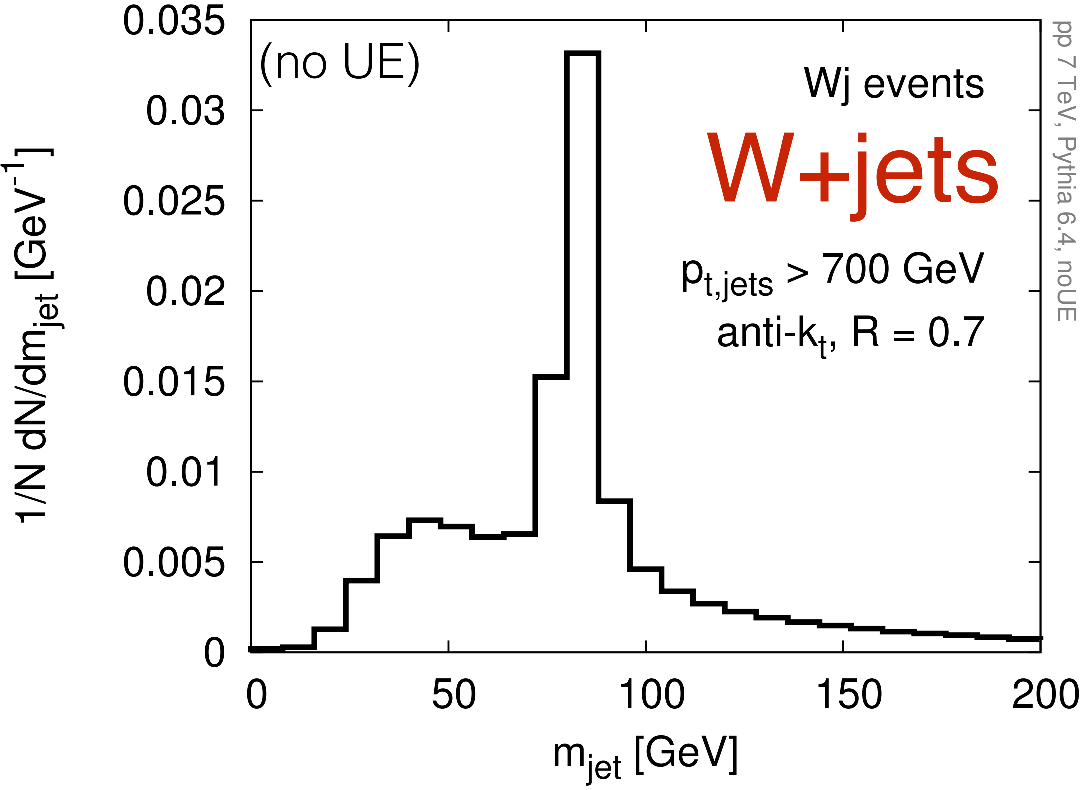
- *Low-mass 4-vector sum of daughters*
- *Many low-angle, asymmetric splittings*



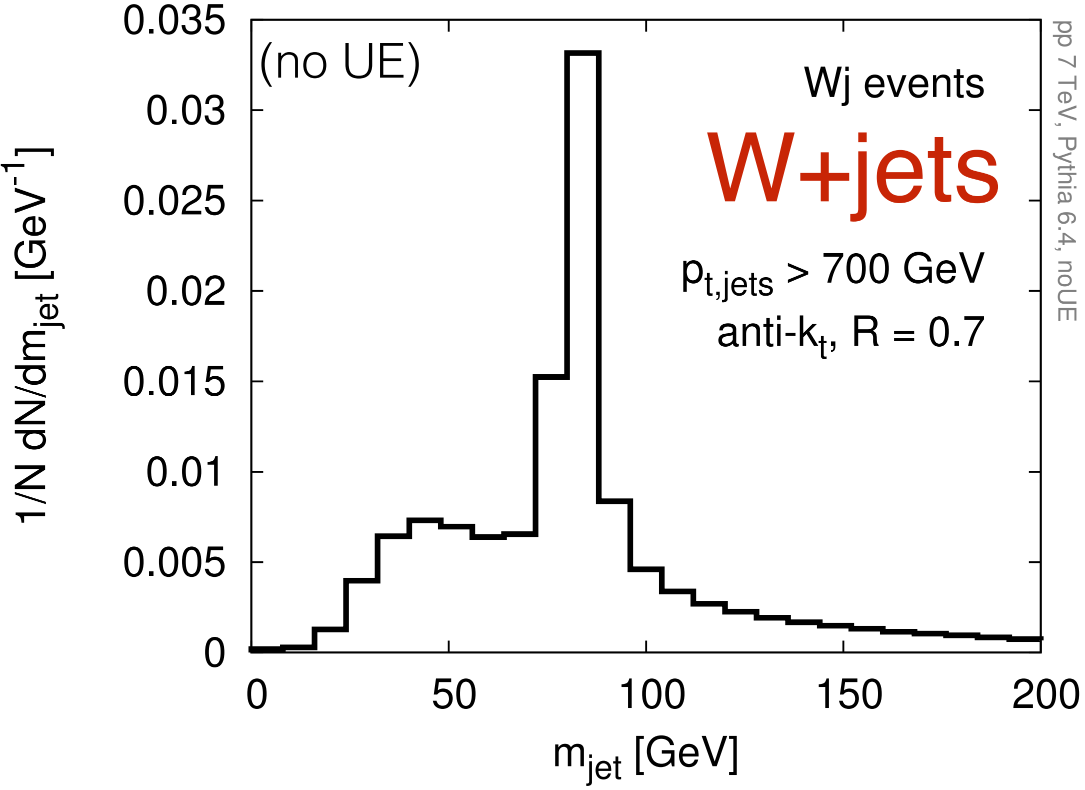
asymmetric splitting

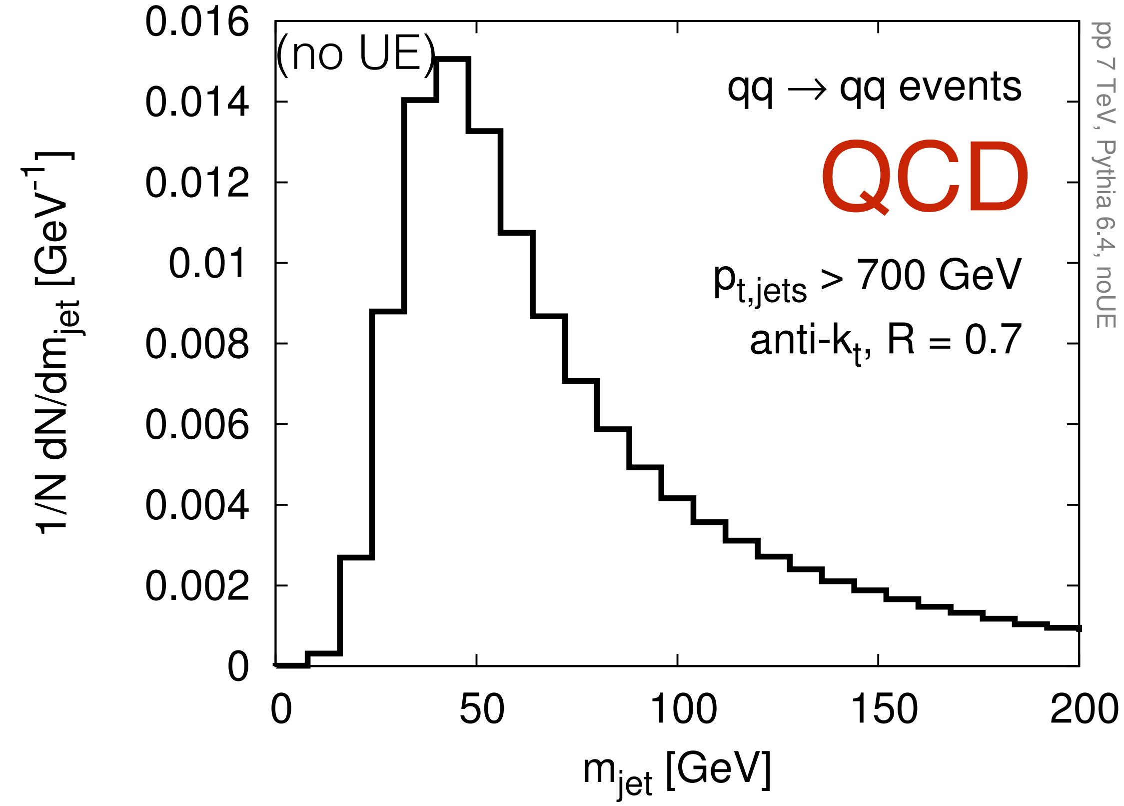
$p_T^q \gg p_T^g$

Jet mass: the problems

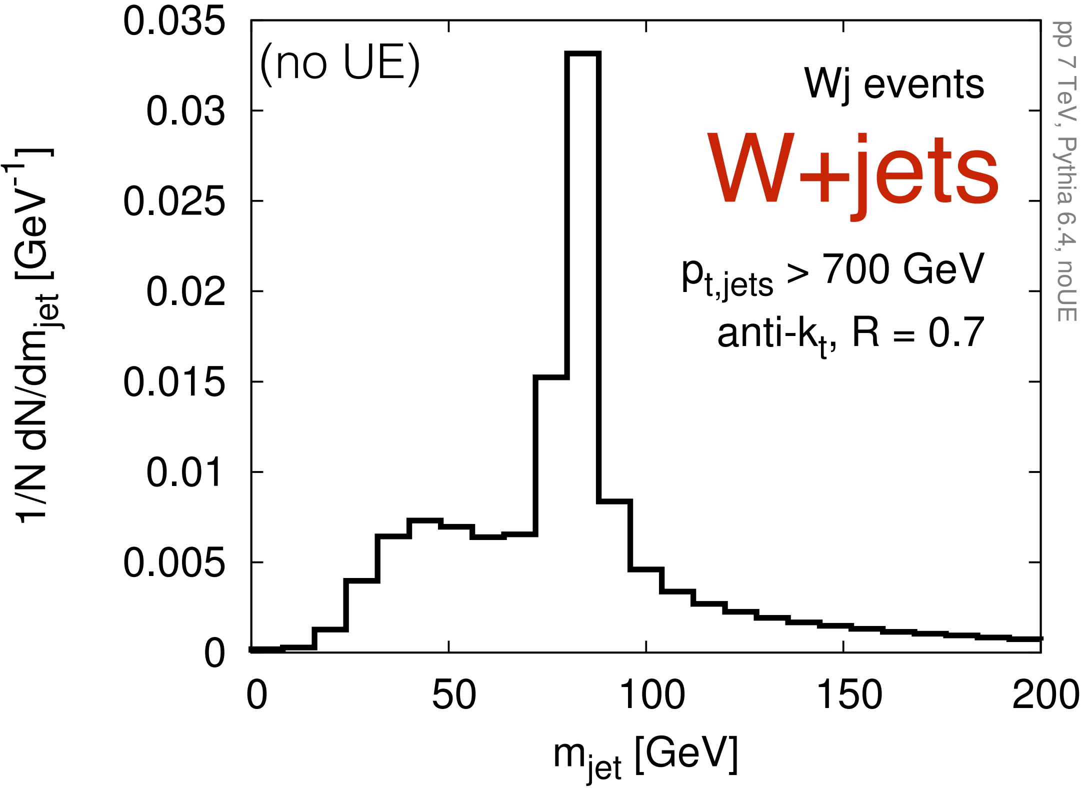


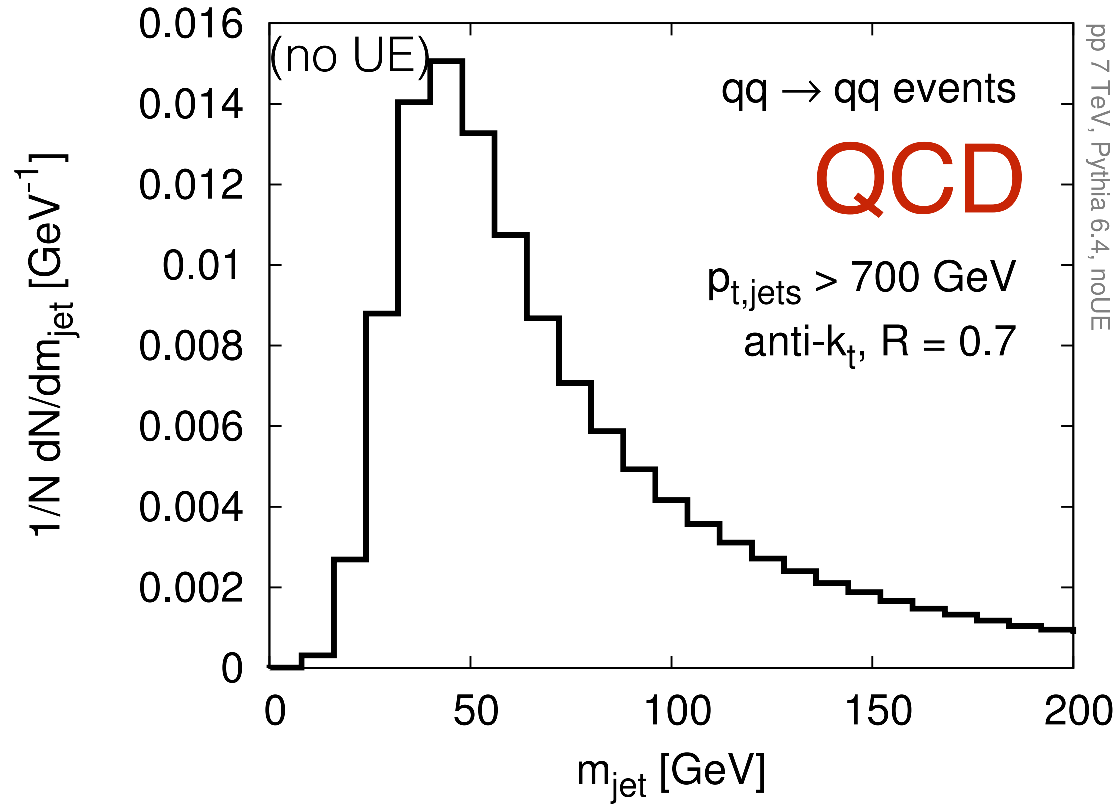
Jet mass: the problems

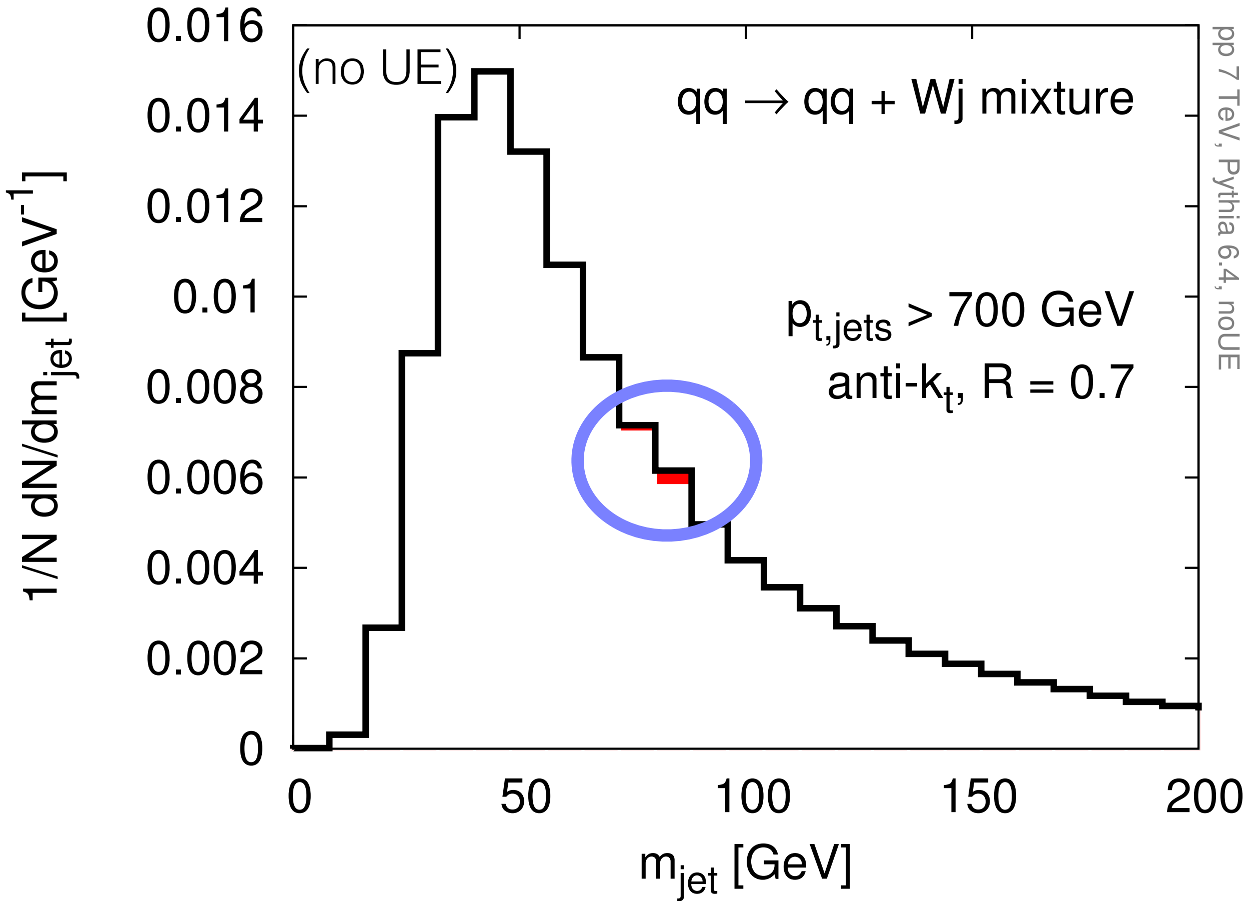




Jet mass: the problems

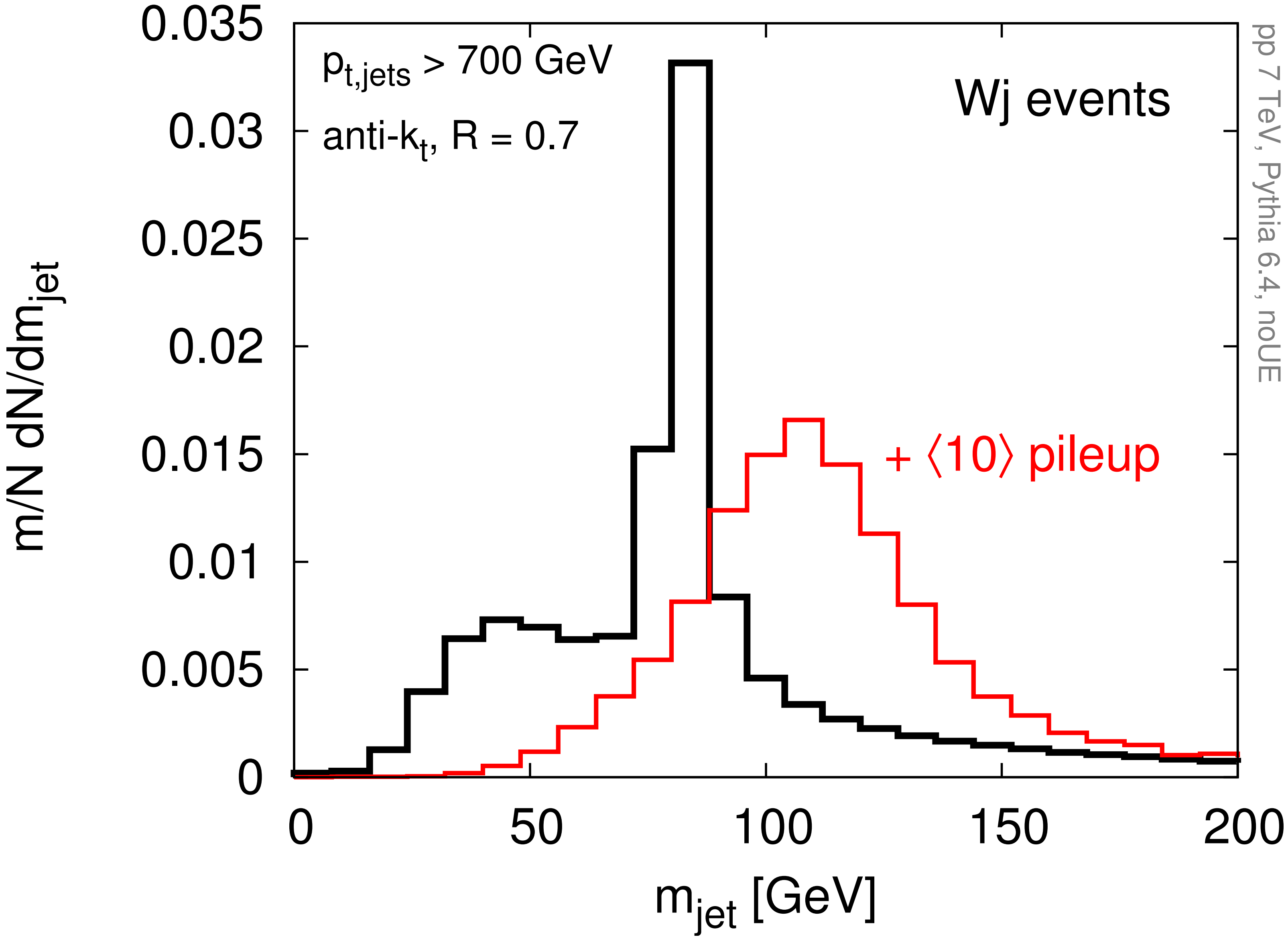






- Have many QCD jets, some of them are massive too!
- HOW TO GET RID OF THEM?

Jet mass: the problems



- Jet mass is EXTREMELY sensitive to pile-up:
  - Jet mass resolution significantly degraded
  - Even more indistinguishable from background!
- Jet substructure hidden by:
  - Underlying event
  - Pileup
  - Soft emissions inside jet

Jet mass: the solution

grooming
/ɡruːmɪŋ/
verb
gerund or present participle of **groom**

To improve signal mass resolution (removing pileup, etc.), without significantly changing background & signal event numbers

- We groom jets in order to remove the softest parts of the jet
- Different jet grooming algorithms :
  - *Trimming*
  - *Pruning*
  - *Softdrop*

Mini recap of jet clustering

Universität Zürich[UZH]

Distance definition:

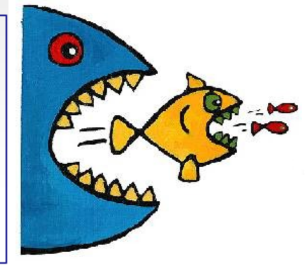
Power of energy vs geometrical scale

$$d_{ij} = \min\left(k_{\mathrm{T}i}^{2p}, k_{\mathrm{T}j}^{2p}\right)\frac{\Delta_{ij}}{D} \qquad \Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

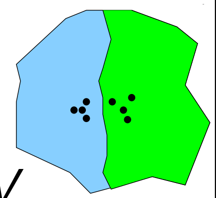
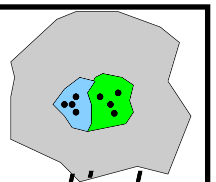
$$d_{iB} = k_{\mathrm{T}i}^{2p}$$

cone size parameter

p = 1: k_T – "Small fish eat first"
p = 0: CA - "Closest fish eat first"
p = -1: anti-k_T "Big fish eat first"

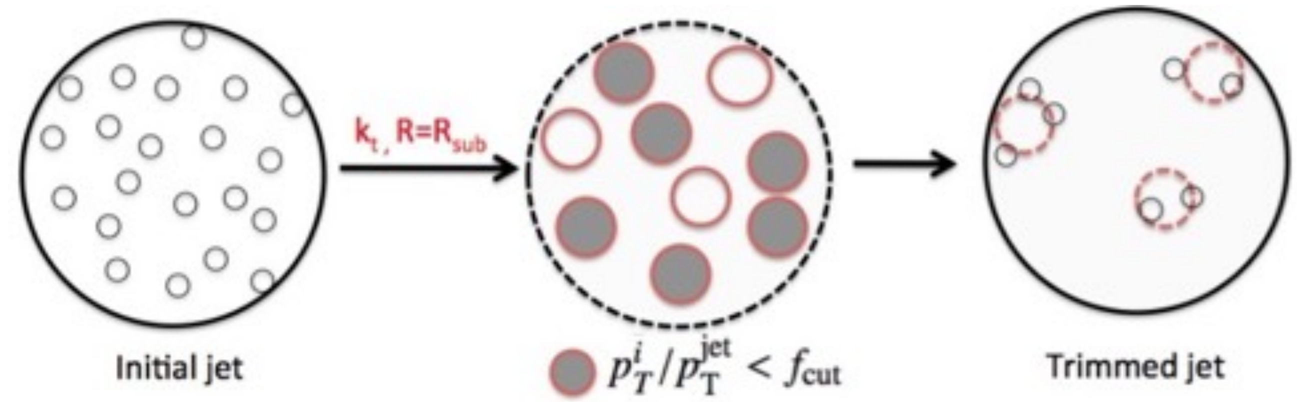


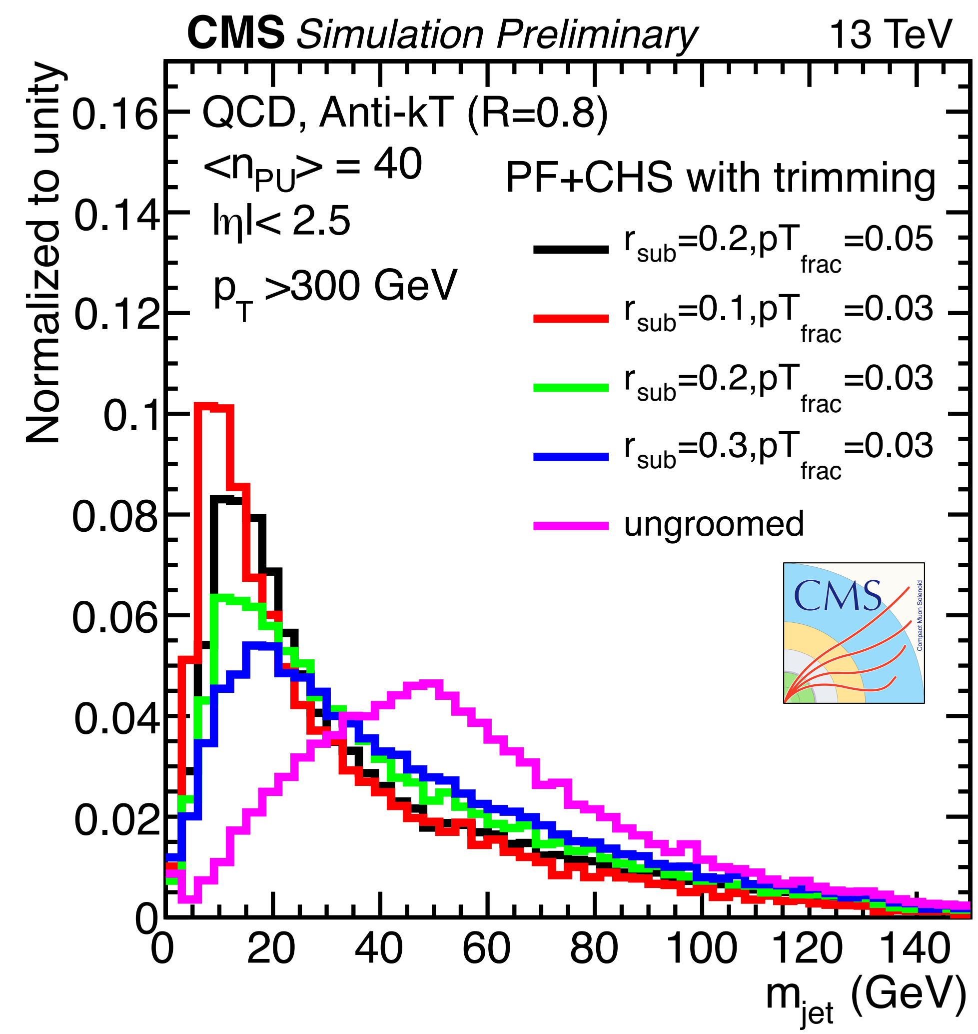
- In CMS, use anti-k_T for jet clustering
- To obtain subjets of jets, undo pair-wise clustering of a jet step-by-step. Need to use different algorithm!

Anti-k_T	k_T 	Cambridge-Aachen 
Useless for substructure *Gradually eats into secondary parton energy deposit* *No clear identification of substructure associated with 2nd parton*	*Clusters soft stuff early* *Last step is merging two hard pieces.* *Easily undone to identify underlying kinematics.* *First algorithm used for jet substructure.*	*Clustering based on spatial separation only* *Shown to provide best performance for resolving jet substructure!* *Pairwise clustering with an angular-ordered structure*

Trimming

- Removes soft constituents from the jet
- Procedure:
  - *Starts with wide jets clustered with anti-k_T or C-A*
  - *Uses k_T algorithm to create subjets of size R_{sub} (with smaller radius R=0.1-0.3) from constituents of the "fat" jet*
  - *Any subjets failing $p_{Ti} / p_T < f_{cut}$ are removed (f_{cut}= 0.03-0.05)*
  - *Assemble remaining subjets into new jet*
- In CMS, you can find trimmed jets used at HLT level as this algorithm is less strict than the other grooming techniques (often used offline)





Tuned parameters: f_{cut} and R_{sub}

Krohn, Thaler, Wang, JHEP 02 (2010) 084

Pruning

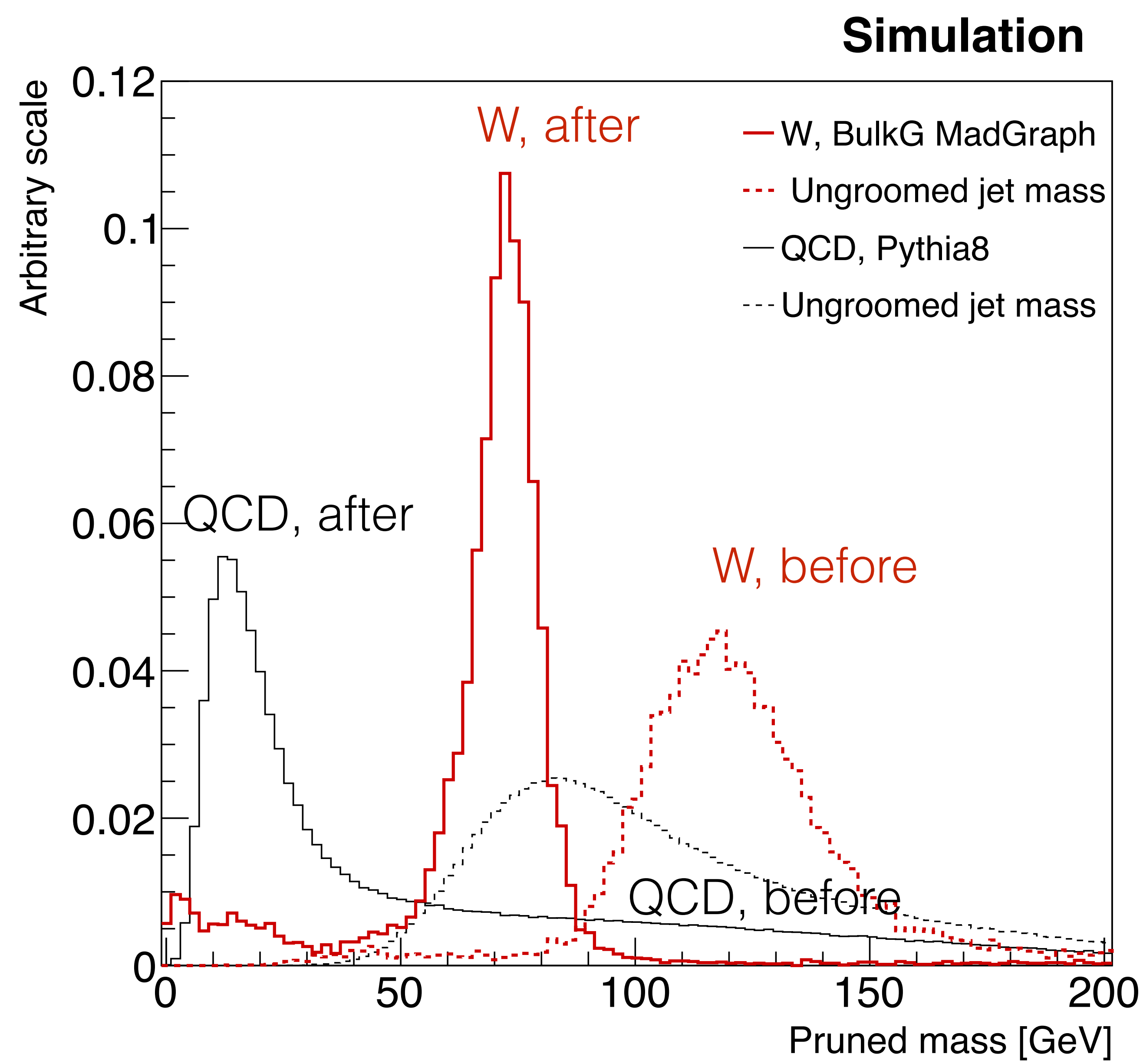
- Removes soft, large angle constituents from the jet
- Procedure:
  - *Recluster jet using Cambridge-Achen algorithm, requiring that each recombination satisfies*

$$\frac{\min(p_{T1}, p_{T2})}{p_{Tp}} > 0.1 \qquad \Delta R_{12} < 0.5 \times \frac{m_{\rm jet}}{p_T}$$

Removes soft particles

Removes wide angle particles

  - *Quark/gluon jet mass pushed to zero!*



Tuned parameters:
z_{cut} and R_{cut}

Ellis, Stephen, Vermilion, Christopher and Walsh
10.1103/PhysRevD.81.094023

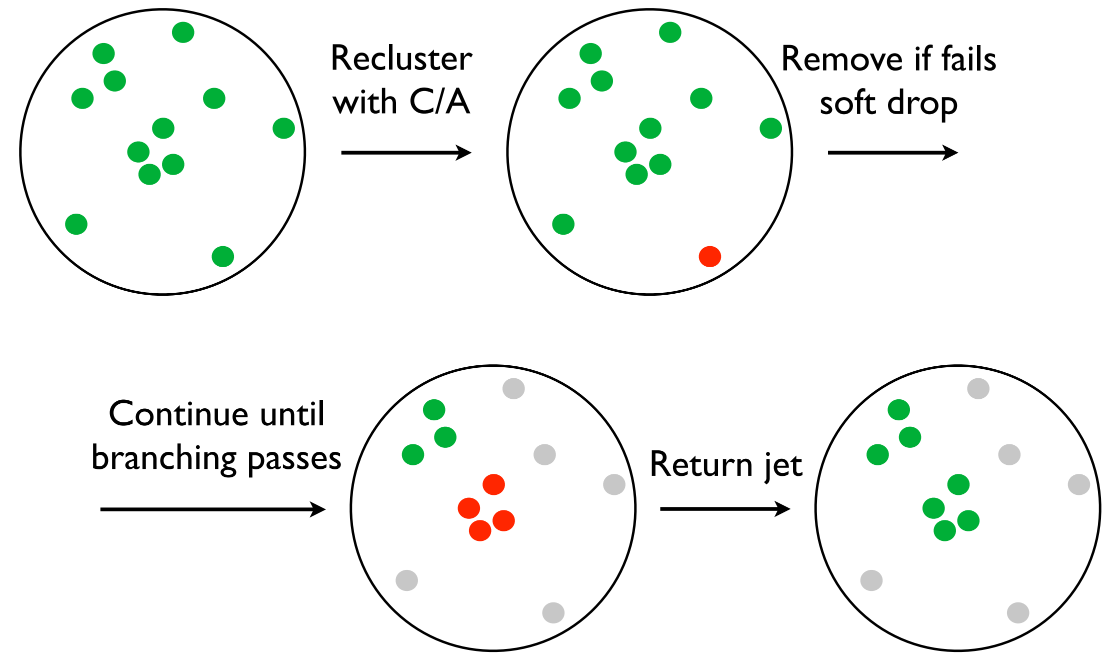
Softdrop

- Procedure:
  - *Recluster jet using C-A algorithm*
  - *Decluster, checking if subjets pass requirements*

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

Soft threshold

Angular exponent



Tuned parameters: z_{cut} and β

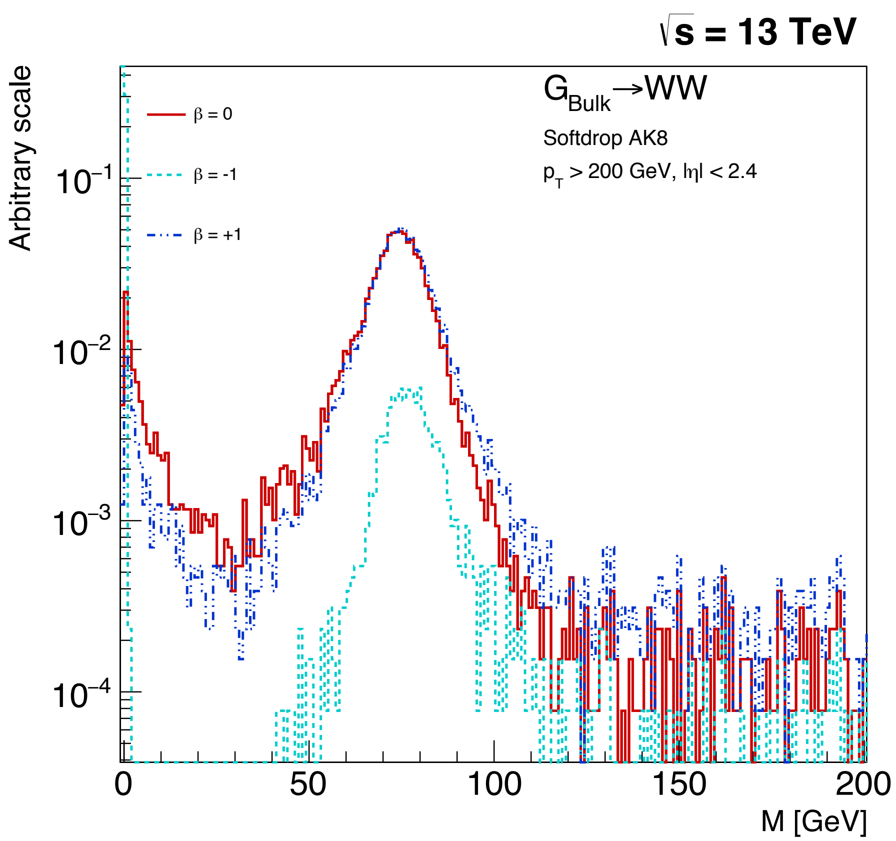
$\beta = \infty$
no grooming

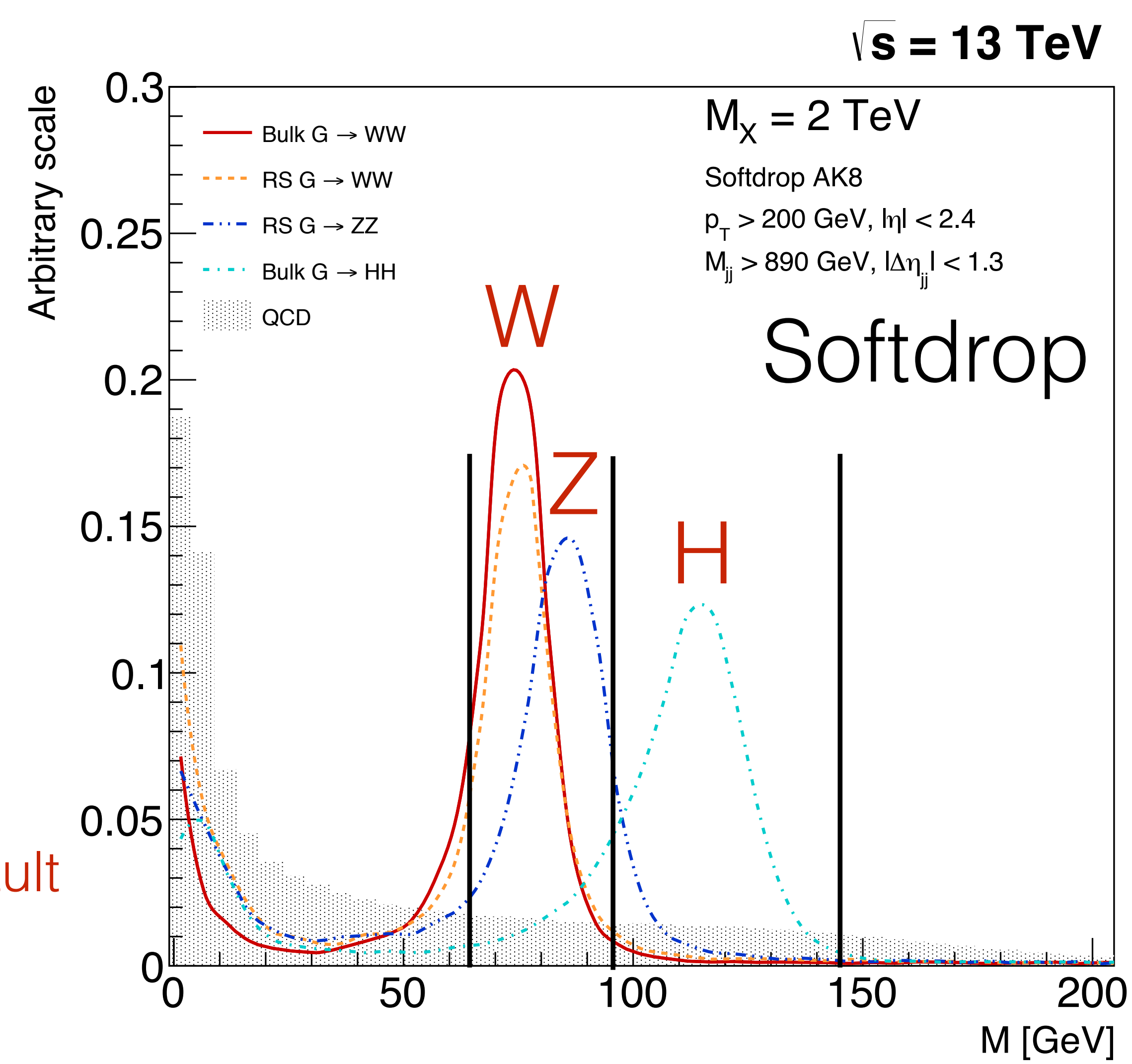
$\beta > 0$
soft, wide angle removed
some soft-collinear removed

$\beta = 0$
all soft emissions removed
modified Mass Drop limit

CMS default

$\beta < 0$
all soft and collinear emissions removed

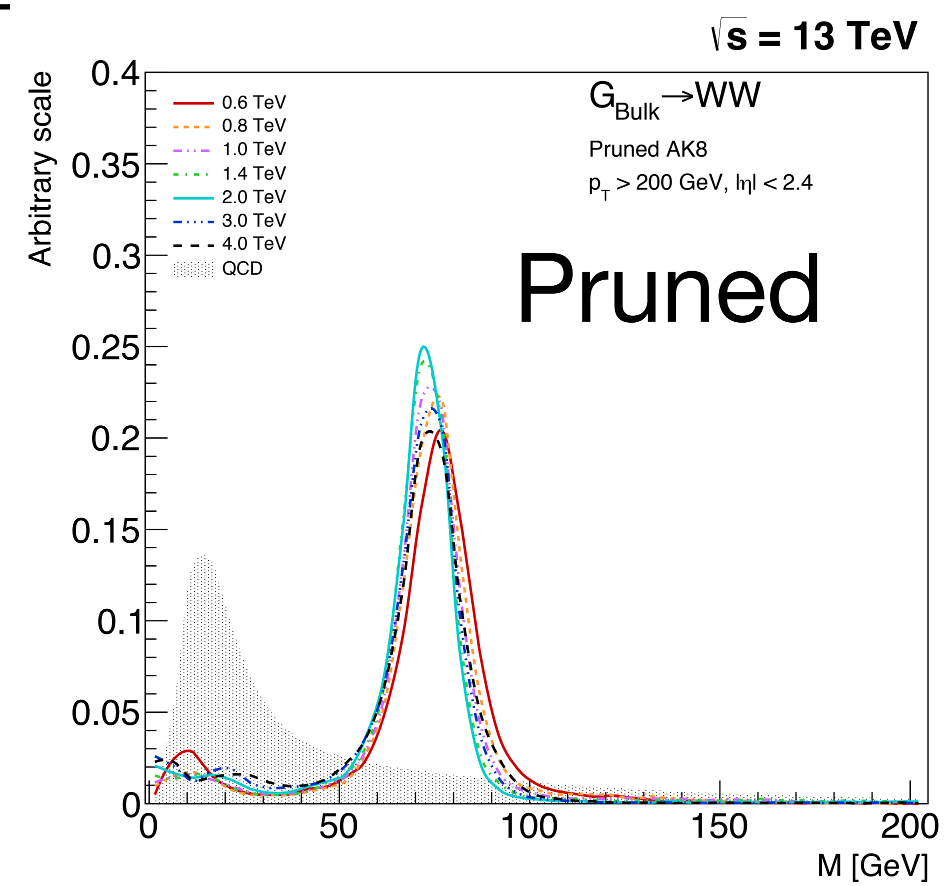


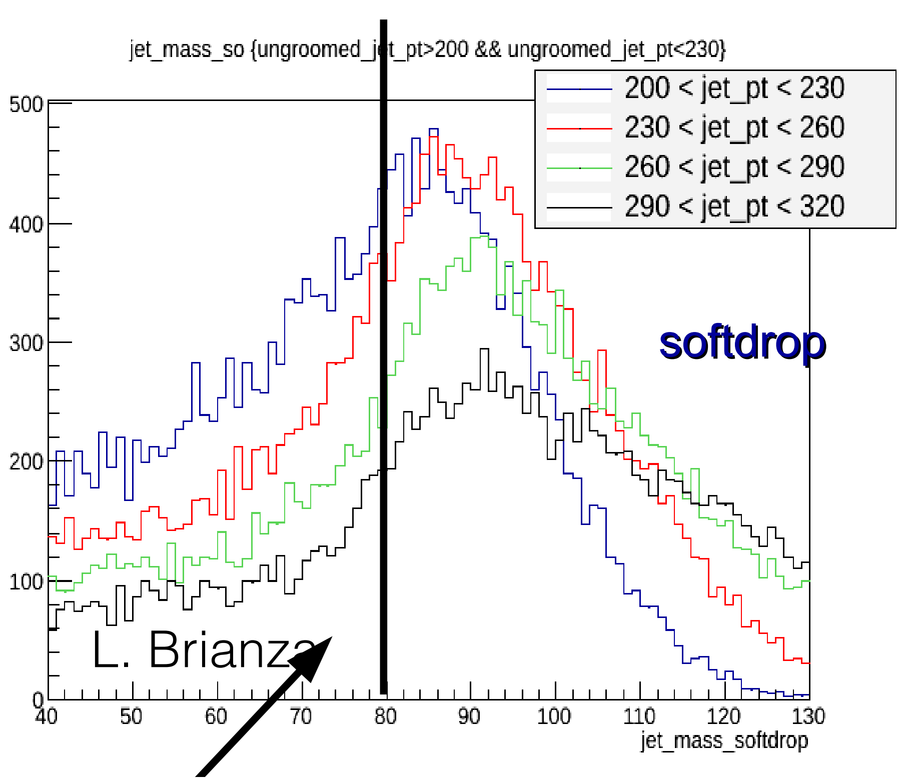


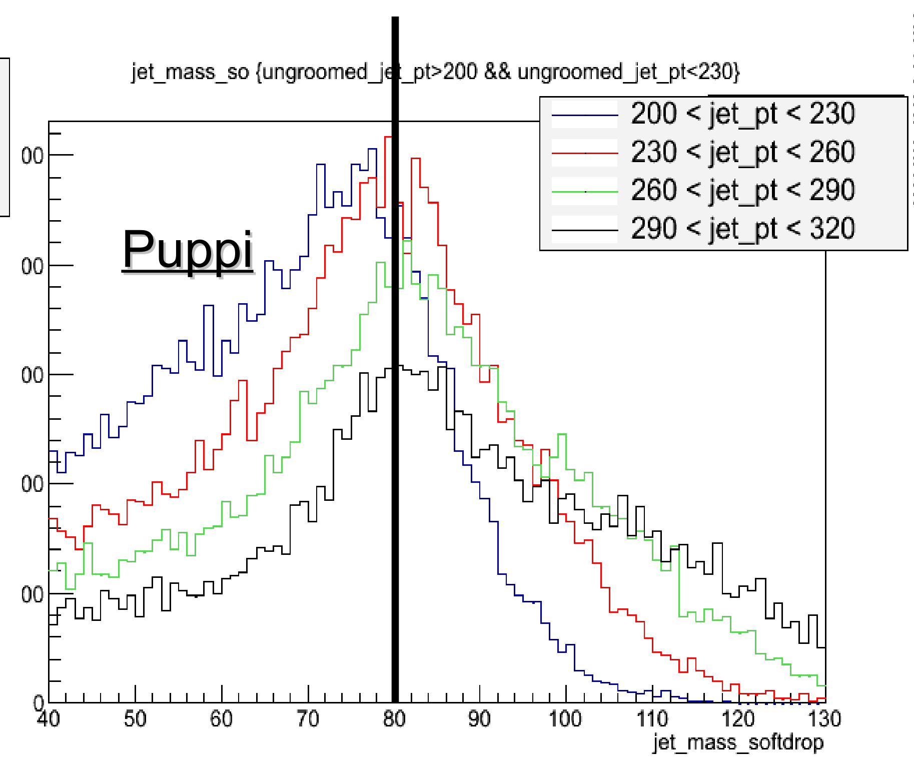
Larkoski, Marzani, Soyez, Thaler, JHEP 05 (2014) 146

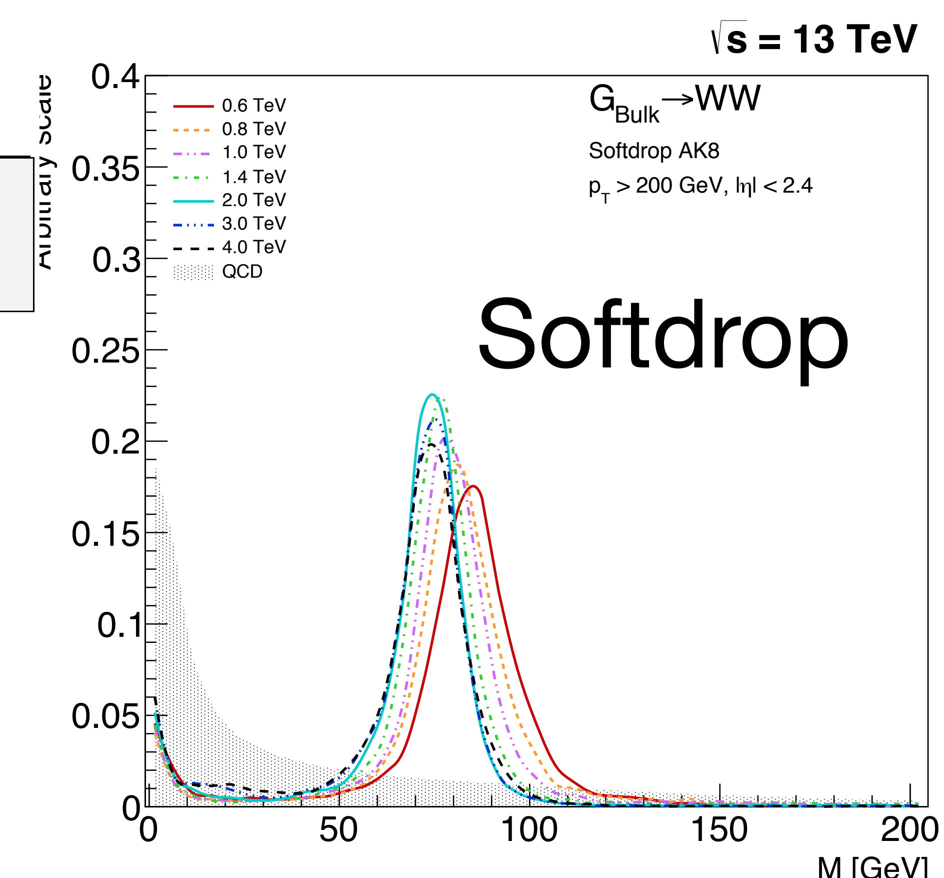
Problems

- p_T-dependent shift observed using softdrop → not optimised to reject underlying event/PU
  - *Peak position can be restored by using soon-to-be-available PUPPI (PileupPerParticle identification) jets*



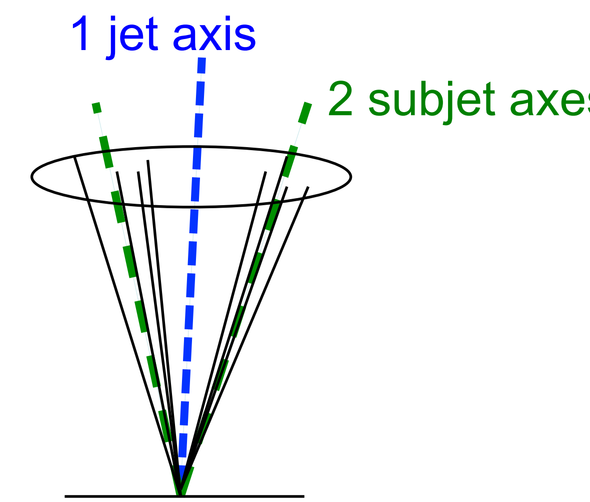






N-subjettiness

- p_T-weighted sum over all jet constituents of the distance w.r.t the closest of N axes in a jet

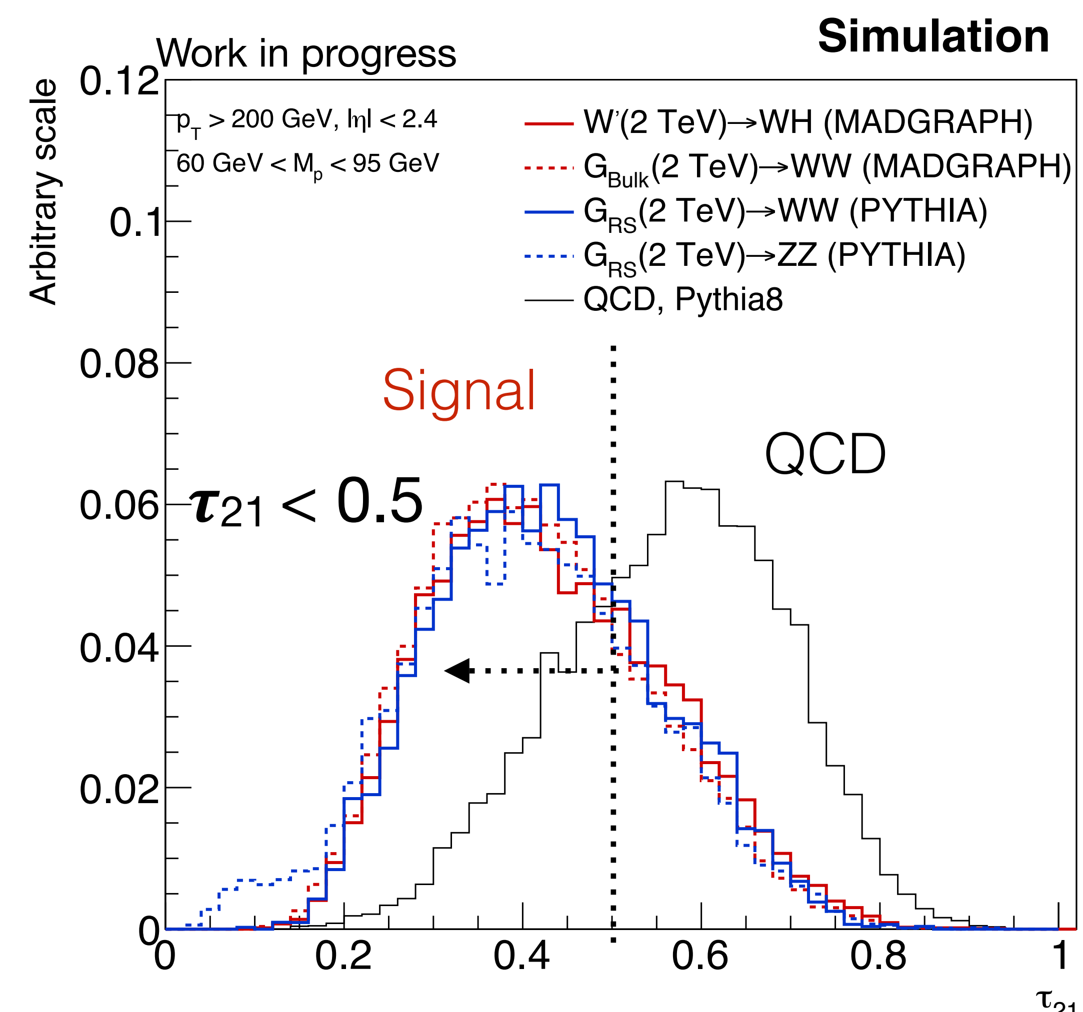


$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\left((\Delta R_{1,k}), (\Delta R_{2,k}) ... (\Delta R_{N,k})\right)$$

Distance between momentum of constituent k w.r.t momentum of rest-frame subjet N

Each constituent assigned to nearest subjet!

  - *Axis are obtained by undoing last (N-1) steps of jet clustering algorithm*
  - *Small τ_N indicates compatibility with N axes hypothesis*
- To discriminate 2-prong W/Z jets from 1-prong q/g jets, use ratio:
  - τ_2/τ_1 (τ_{21})
- To discriminate 3-prong top jets from 1-prong or 2 prong jets, use ratio:
  - τ_3/τ_1 (τ_{31}) or
  - τ_3/τ_2 (τ_{32})



Jet substructure at CMS

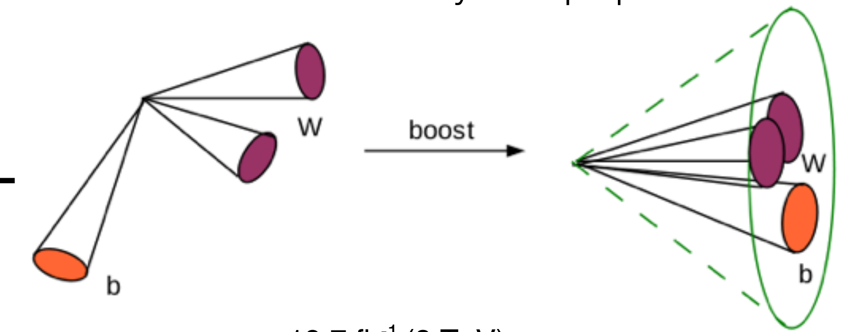
tagging
/taggɪŋ/
verb
gerund or present participle of **tag**

Reducing the background as much as you can, leaving much of the signal

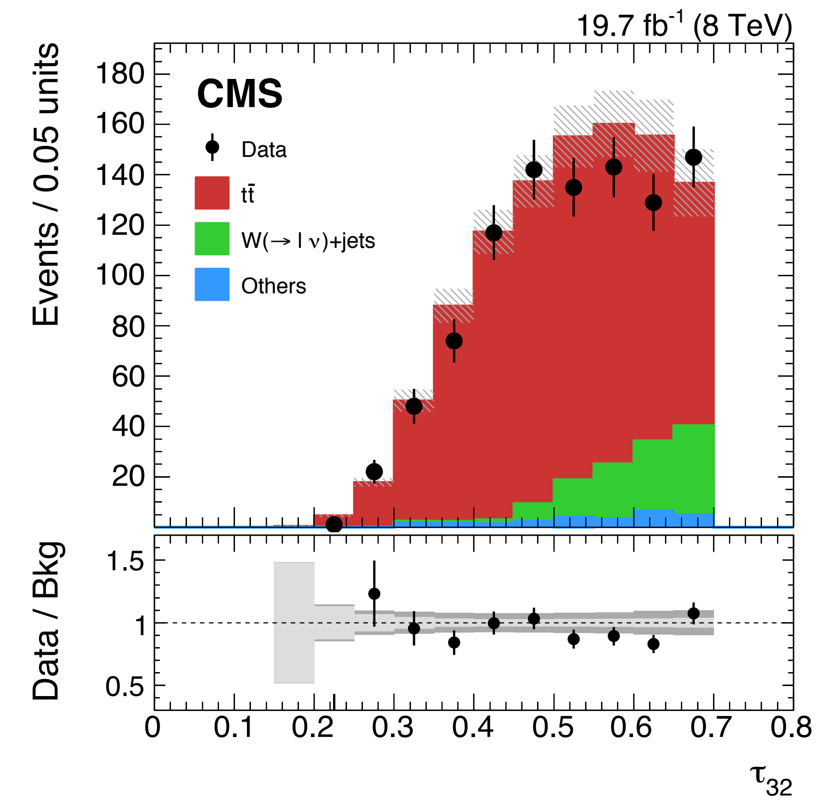
- At CMS, combine the methods mentioned previously in order to "tag" jets
- Different jet tagging algorithms :
  - *Top tagger*
  - *W/Z tagger*
  - *(H tagger)*

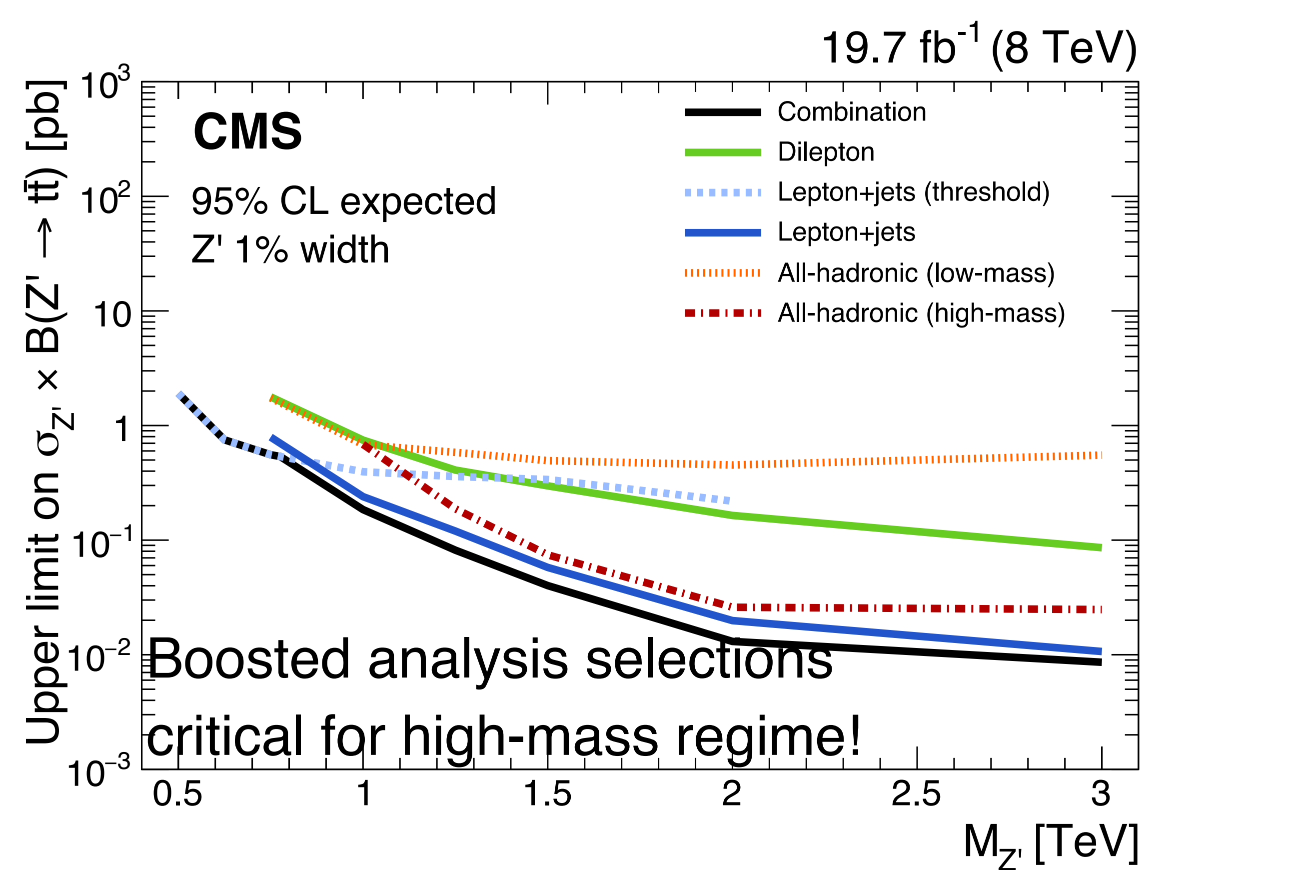
Top tagging

Hadronic decay of top quark



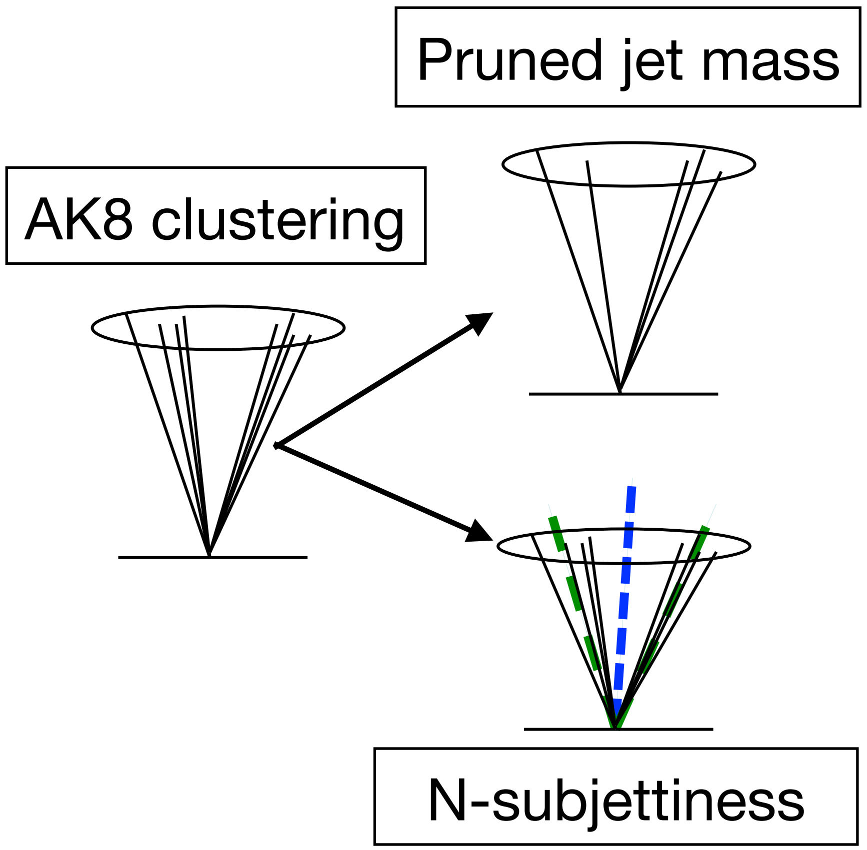
- Two algorithms in use at CMS
  - *HEPTopTagger algorithm (For lower boosts, use R = 1.5 jets)*
  - *CMSTopTagger algorithm (For high boosts, R = 0.8 jets. Better efficiency for p_T above ~400 GeV)*
- CMS top tagger:
  - *Jet mass in [140, 250] GeV window*
  - *3 or more subjets*
  - *Minimum di-subjet mass above 50 GeV*
  - *N-subjettiness $\tau_{32} < 0.7$*

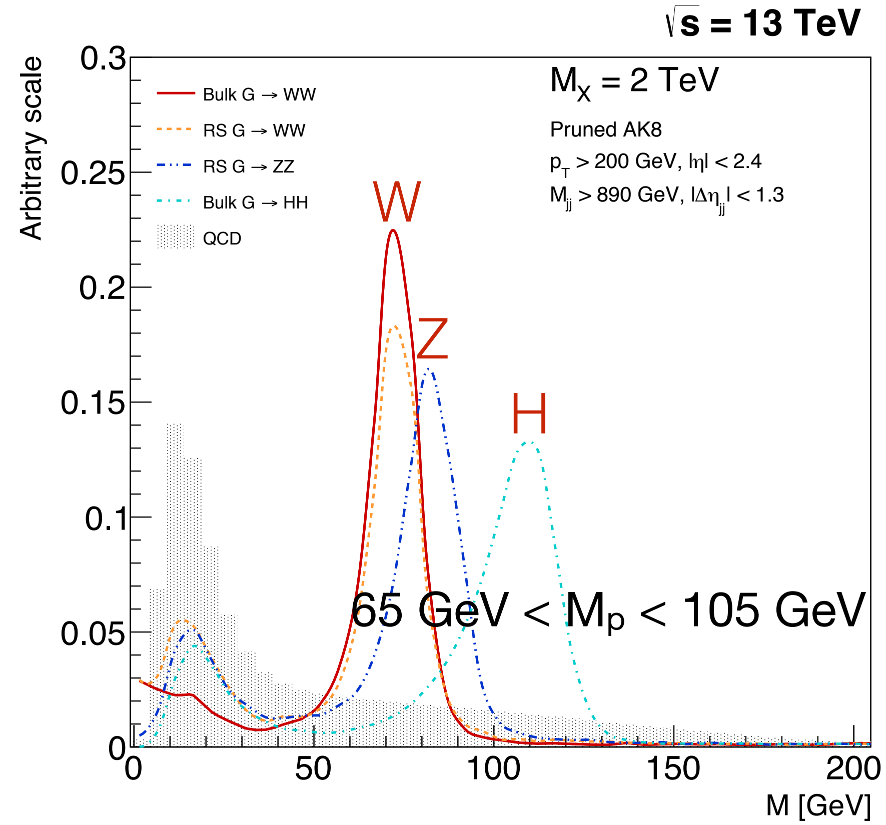


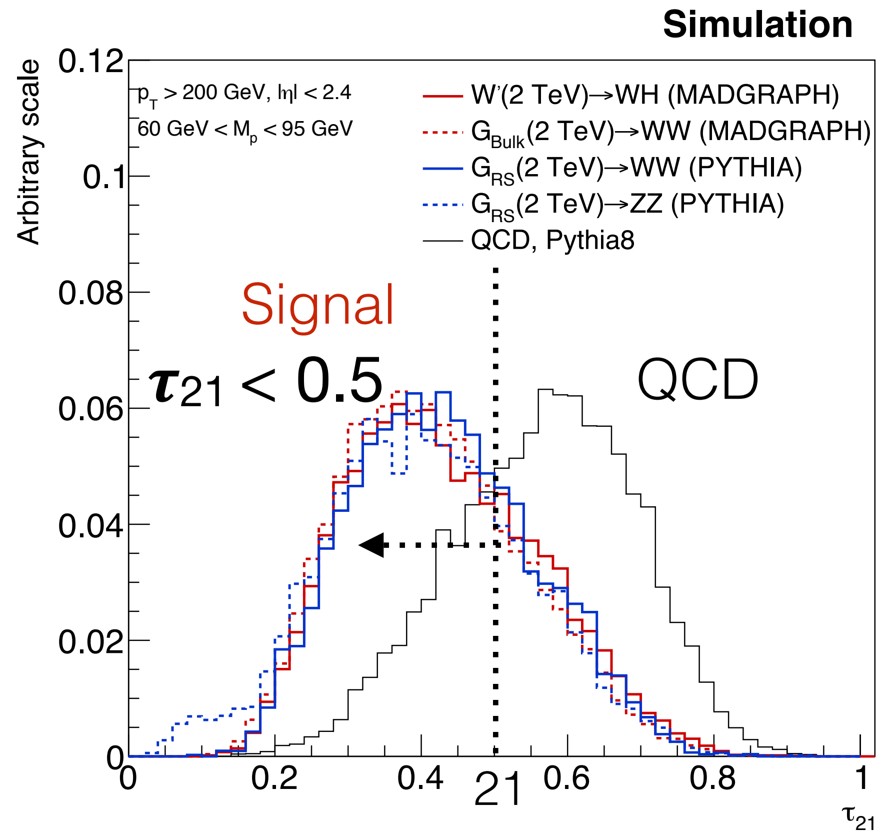


Boosted analysis selections critical for high-mass regime!

W/Z-tagging





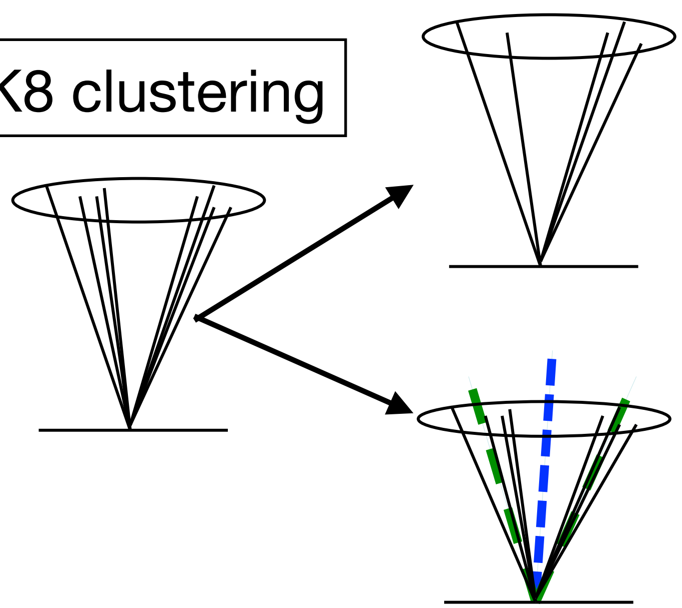


Typical W/Z-tagger (pruning+n-subjettiness):
- 3 % mistag rate
- 55 % efficiency

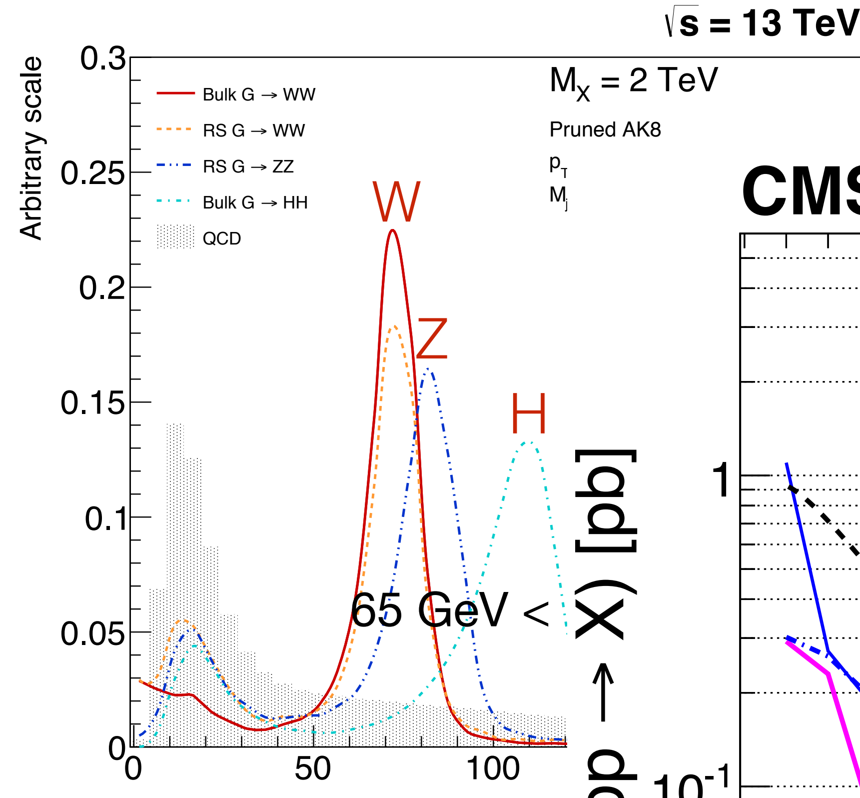
W/Z-tagging

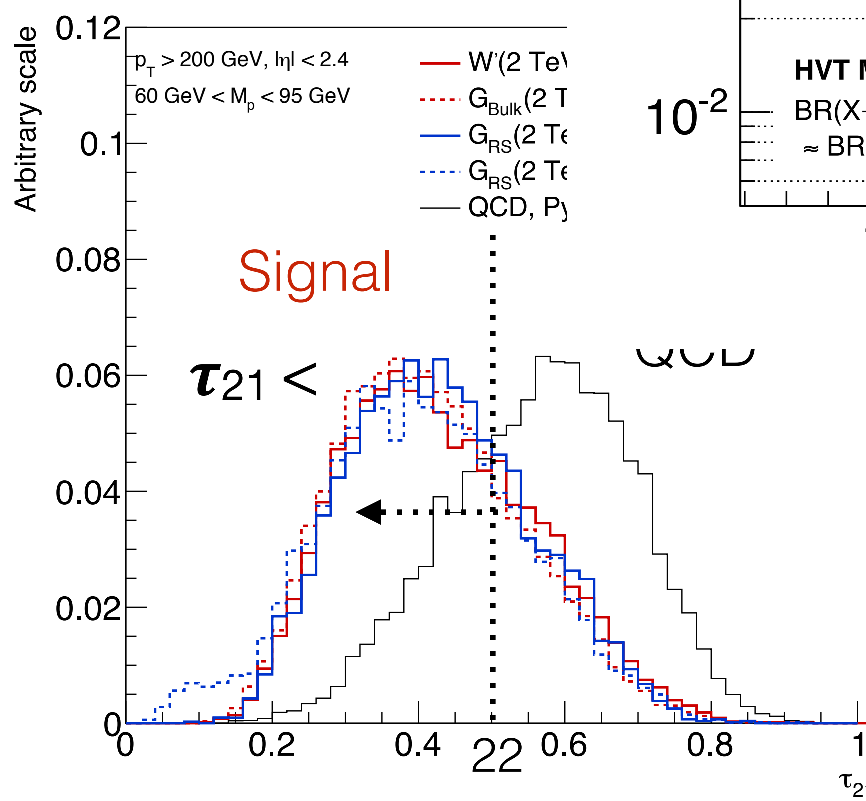
Pruned jet mass

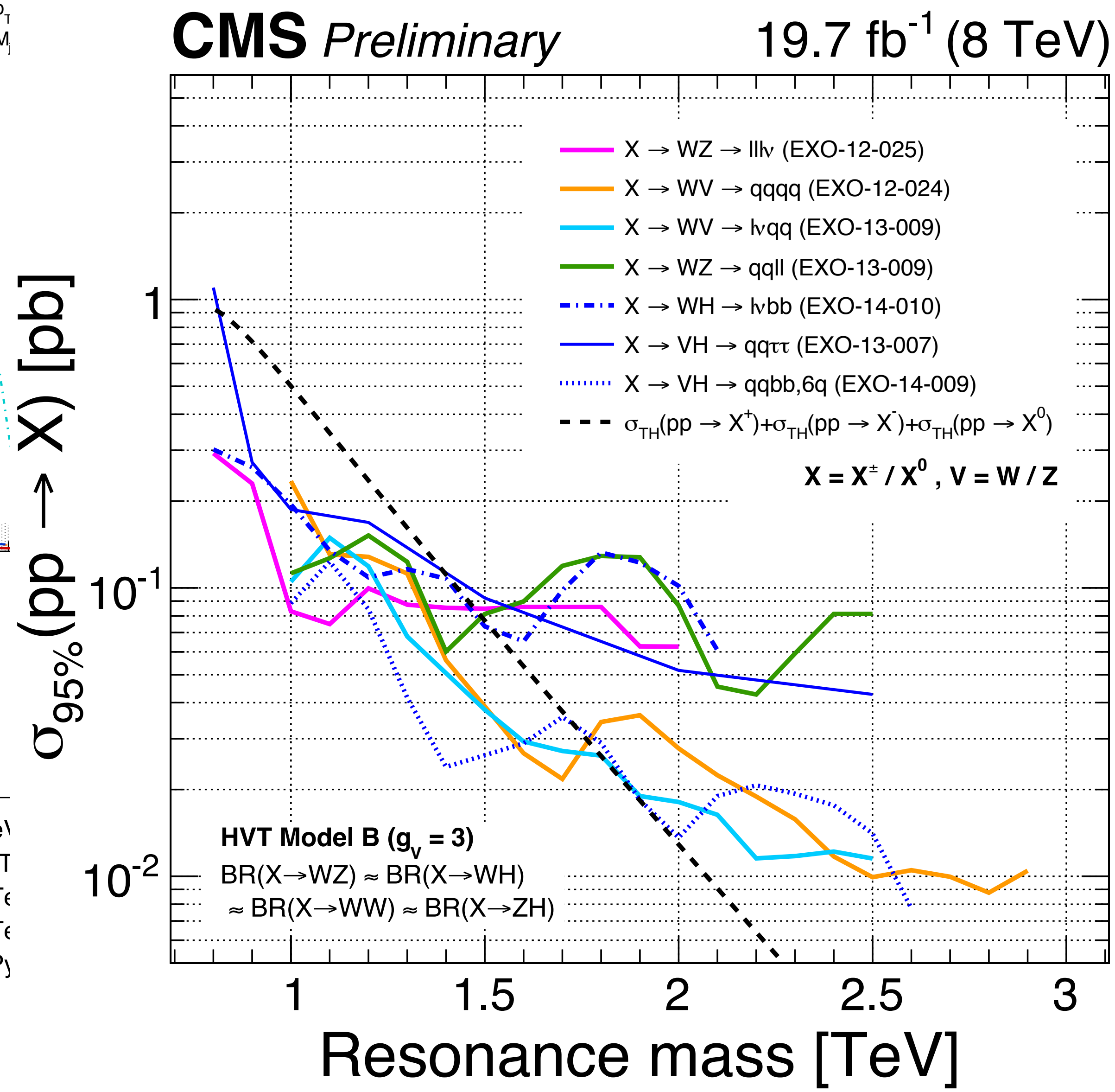
AK8 clustering



N-subjettiness







EXO-12-024
EXO-13-009
arXiv:1506.01443
CMS-PAS-EXO-14-010
CMS-PAS-EXO-14-009
CMS-PAS-EXO-14-006
....

Summary

Universität Zürich[UZH]

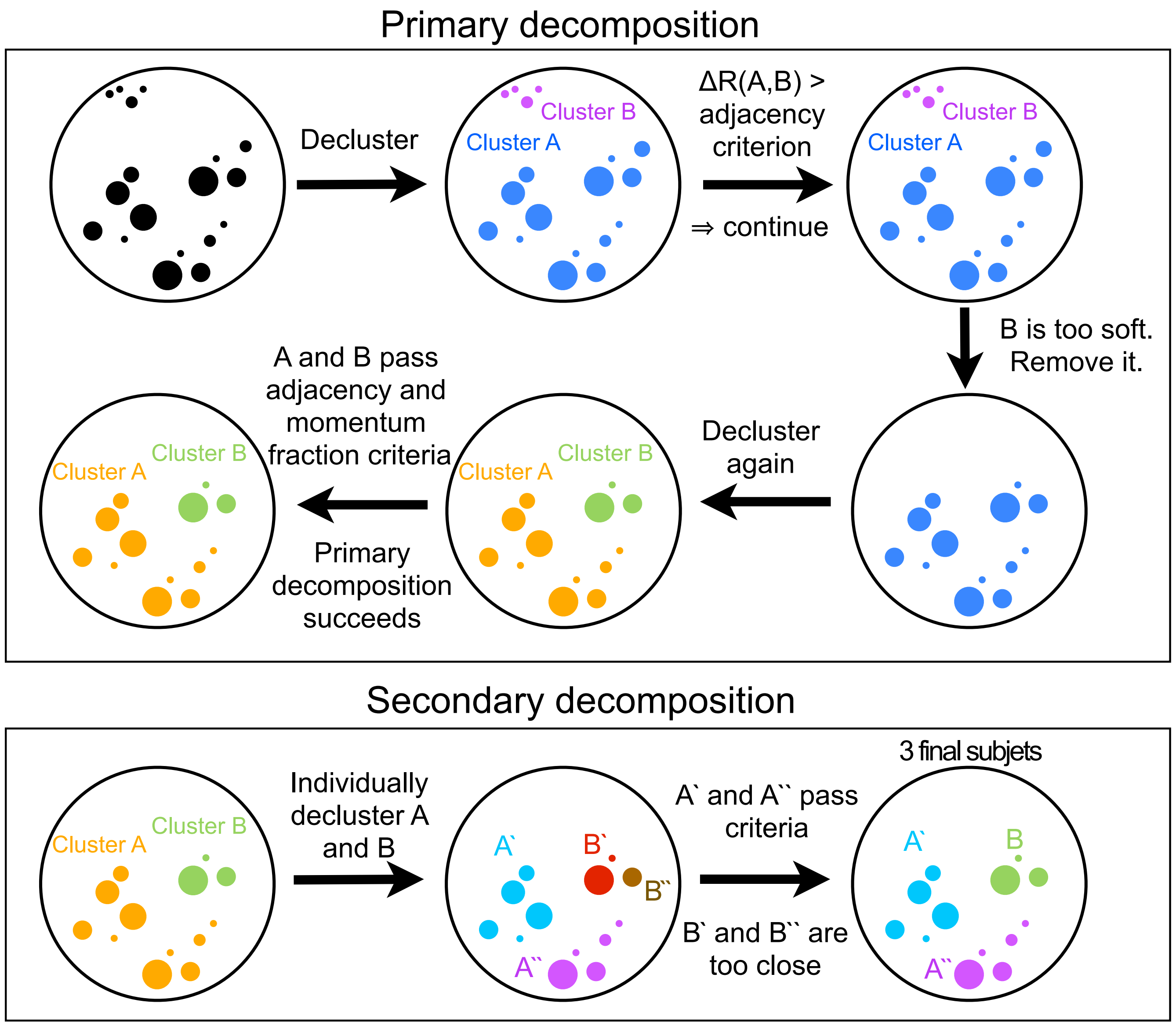
- Use of novel W/Z tagging techniques allows great discrimination between signal and background in boosted topologies
  - *Very important tool to maintain sensitivity at 13 TeV LHC!*
- Many analyses at CMS are currently using boosted top and W/Z taggers. Fast evolving field!
- With these new methods we will be able to exclude (discover!) particles with masses in the 1-4 TeV range. Perhaps already this year?!

Universität Zürich[UZH]

BACKUP

CMS Top Tagger

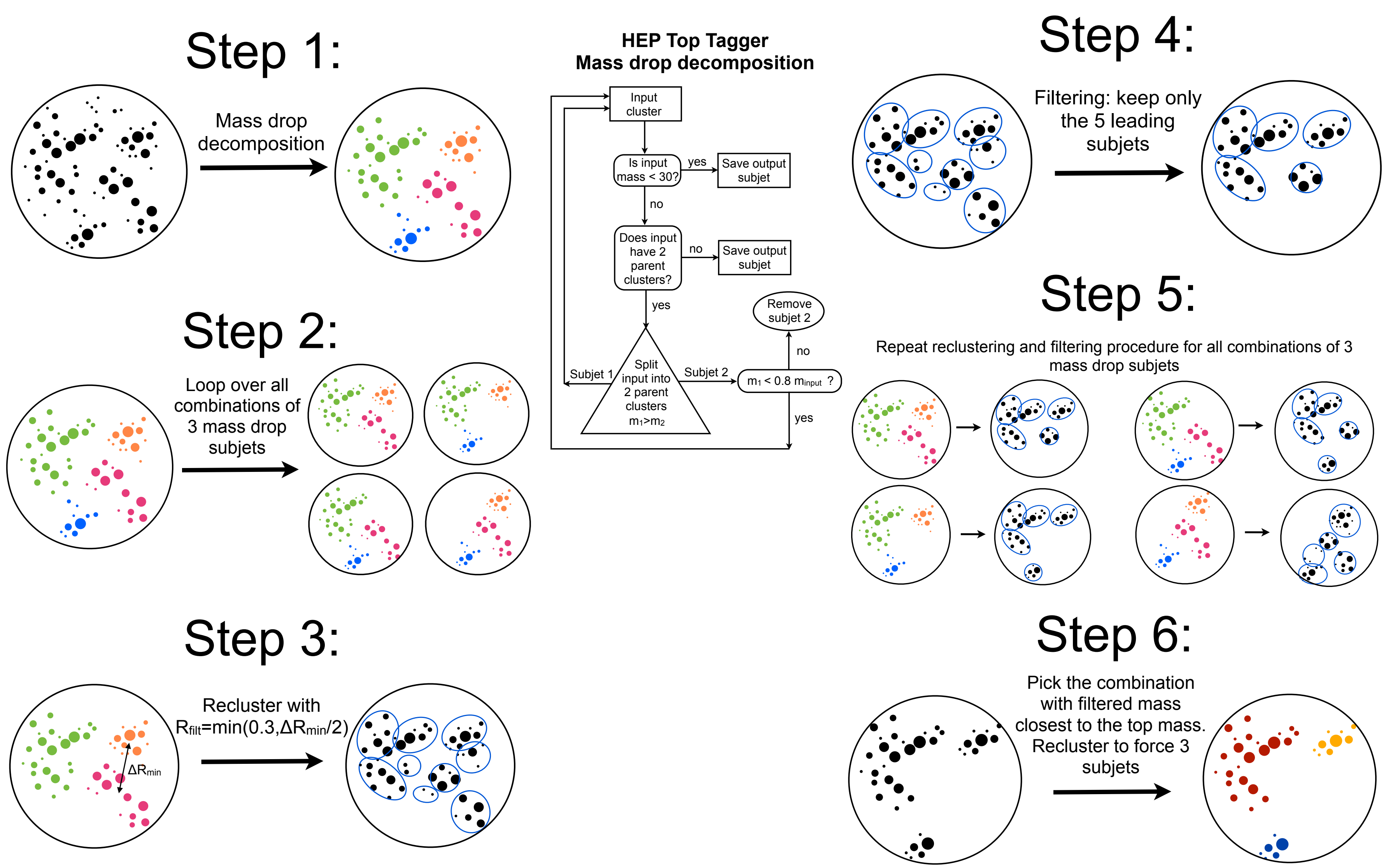
Example: CMS Top Tagger decomposition



38

HEP Top Tagger

HEP Top Tagger details



39