Semiclassical approach to jet clustering and background subtraction

Jeff Tseng
in collaboration with
Hannah Evans, Jesse Liu, and Hongbin Chen

13 August 2013

5th International Joint Theory/Experiment Workshop on
BOOSTED OBJECT PHENOMENOLOGY, RECONSTRUCTION & SEARCHES IN HIGH ENERGY COLLISION EXPERIMENTS

Motivation / Outline

- Work grew out of trying myself to learn more about jet clustering
  - In particular, non-deterministic methods like Qjets
  - Different ways to turn probabilities into jets
- ScJet arose from thinking about classical radiation
- “ScSubJet”: reformulated for higher pileup levels
- Compare with existing techniques in highly idealized (nearly toy) simulation tests
- “We don't need theorists to convince experimentalists to do stupid things” - J Thaler, 12 Aug 2013
  - How naive can one be and still end up with jets?

ScJet

arXiv:1304.1025
PRD 88, 014044 (2013)

- Sequential recombination algorithm
- Inter-cluster distance

$$d_{ij} = \frac{1}{4}\left(m_{Ti} + m_{Tj}\right)^2 \left(\frac{\Delta R_{ij}}{R}\right)^3$$

$$m_{Ti}^2 = m_i^2 + p_{Ti}^2$$

$$\Delta R_{ij}^2 = \Delta \varphi_{ij}^2 + \Delta y_{ij}^2$$

  - R = maximum ΔR_{ij} for merging

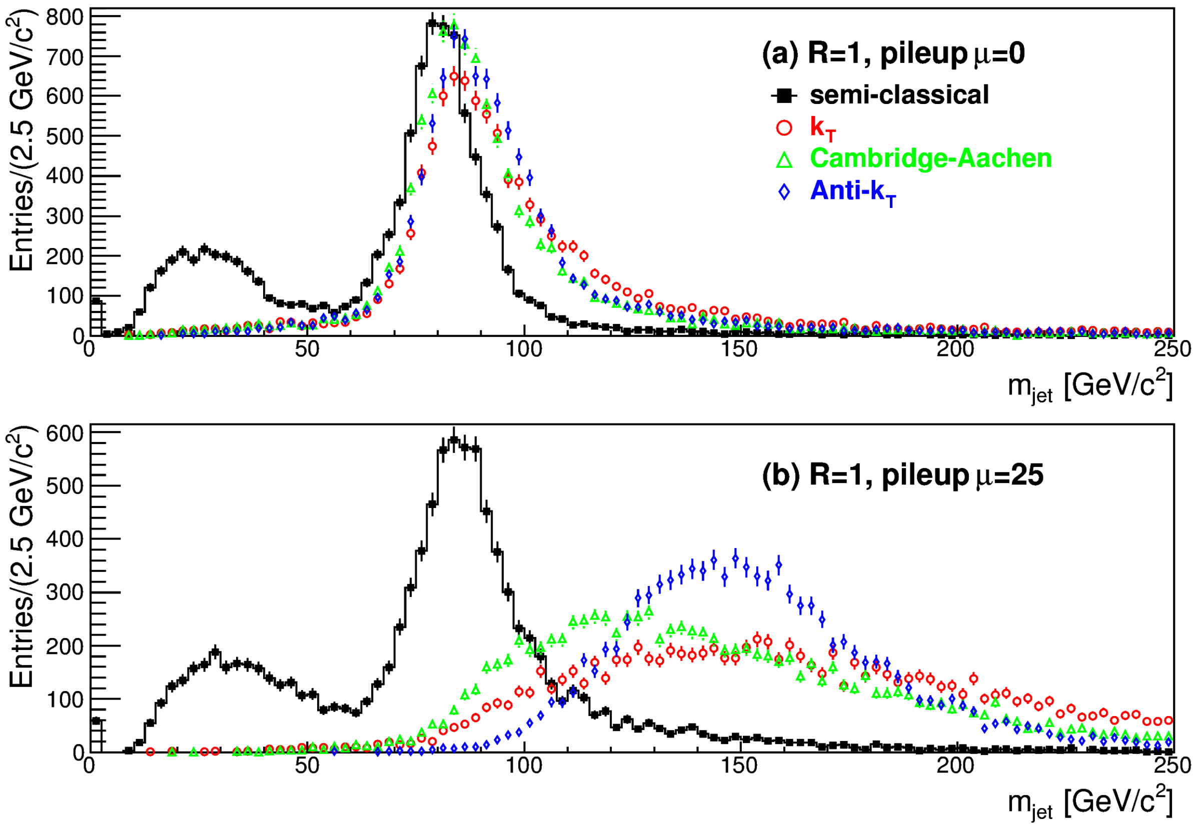
- Beam-cluster distance $d_{iB} = m_{Ti}^2$

- Merge clusters by adding 4-momenta

Simulation tests

- Pythia 8.176 generation, 8 or 14 TeV cm energy
  - W+jet, p_T > 500 GeV
  - Minbias pileup with Tune 4Cx
- Toy “detector”
  - Cluster particles into 0.1x0.1 ΔφxΔη towers
  - Remove ν and charged particles p_T < 0.4 GeV
- Fastjet 3.0.4 for clustering, grooming
  - ScJet 1.1.0 plugin available from Fastjet contrib website http://fastjet.hepforge.org/contrib/ (fjcontrib ≥ 1.005)

Comparison



- Ungroomed jet clustering:
  - k_T
  - Cambridge-Aachen
  - Anti-k_T
- 0 and 25 average pileup

ScJet and pruning

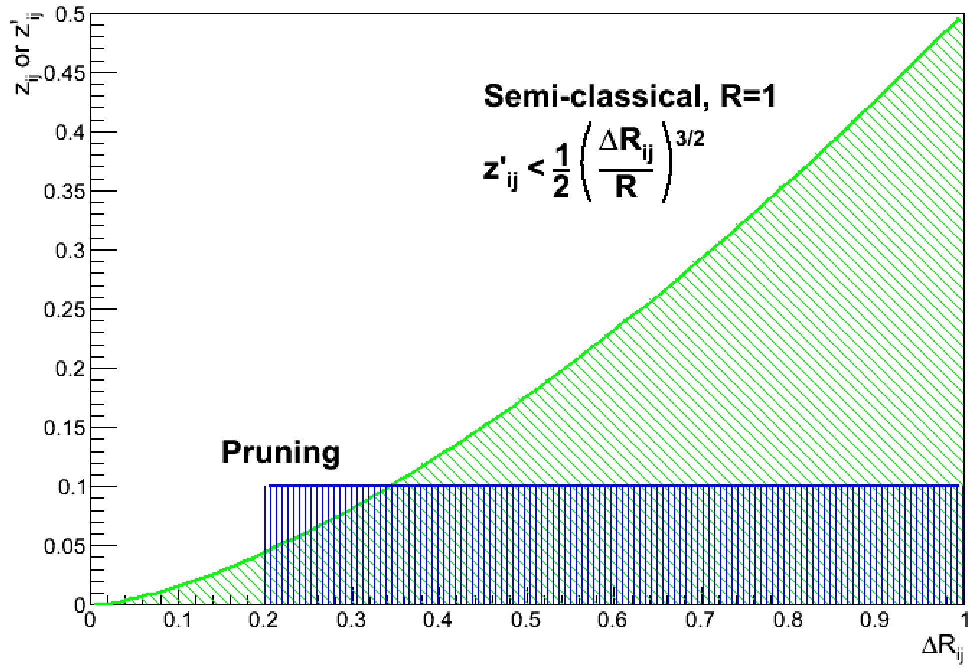
- ScJet eliminates clusters while clustering

$$z'_{ij} \equiv \frac{m_{Ti}}{m_{Ti}+m_{Tj}} < \frac{1}{2}\left(\frac{\Delta R_{ij}}{R}\right)^{3/2}$$

- Similar to pruning

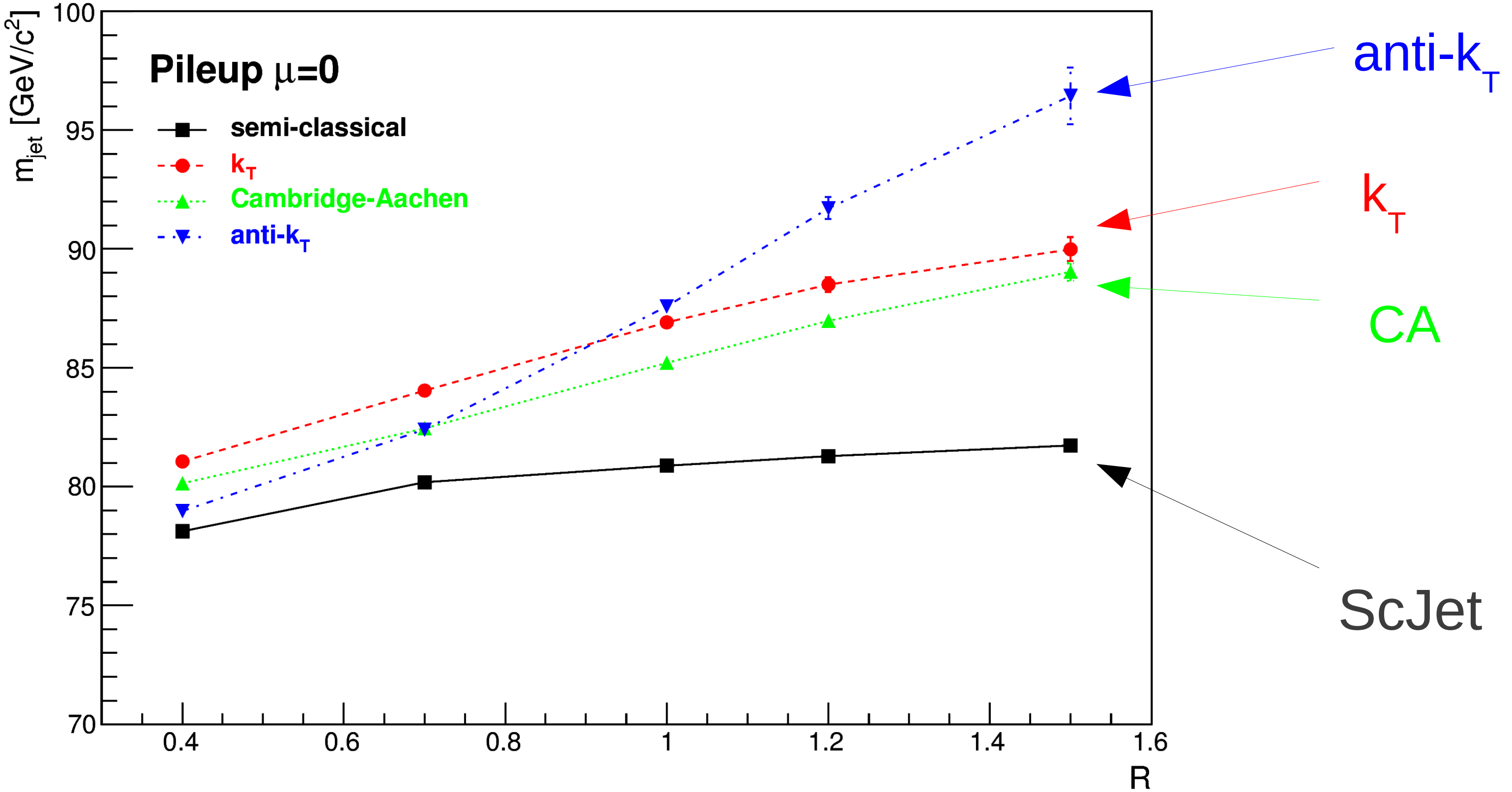
$$z_{ij} \equiv \frac{min(p_{Ti}, p_{Tj})}{|\vec{p}_{Ti}+\vec{p}_{Tj}|} < z_{cut}$$

$$\Delta R_{ij} > D_{cut}$$



- Not surprising if there is some similar behavior
- Jet area identically zero

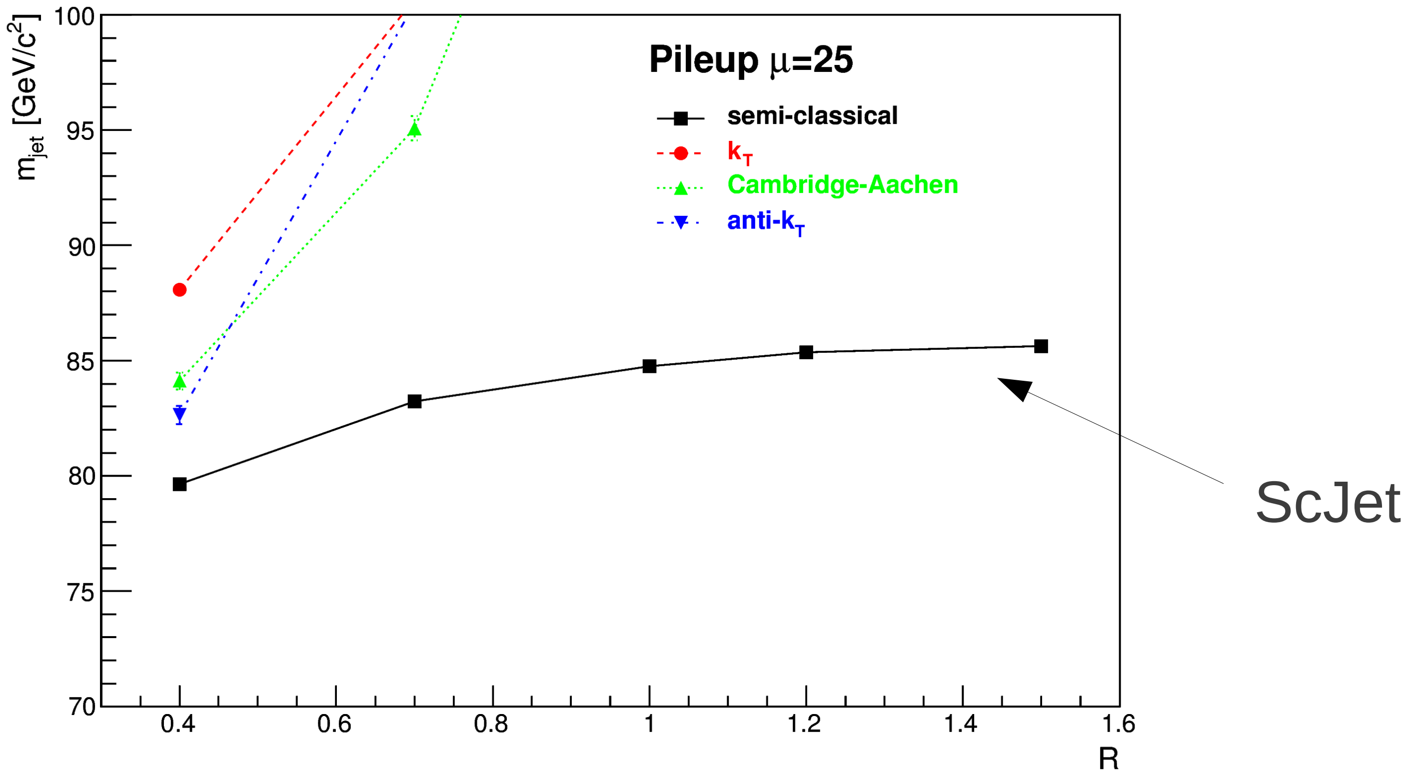
Stability vs jet size



- W peak mass variation with *R*
- ScJet appears to level off with larger *R*

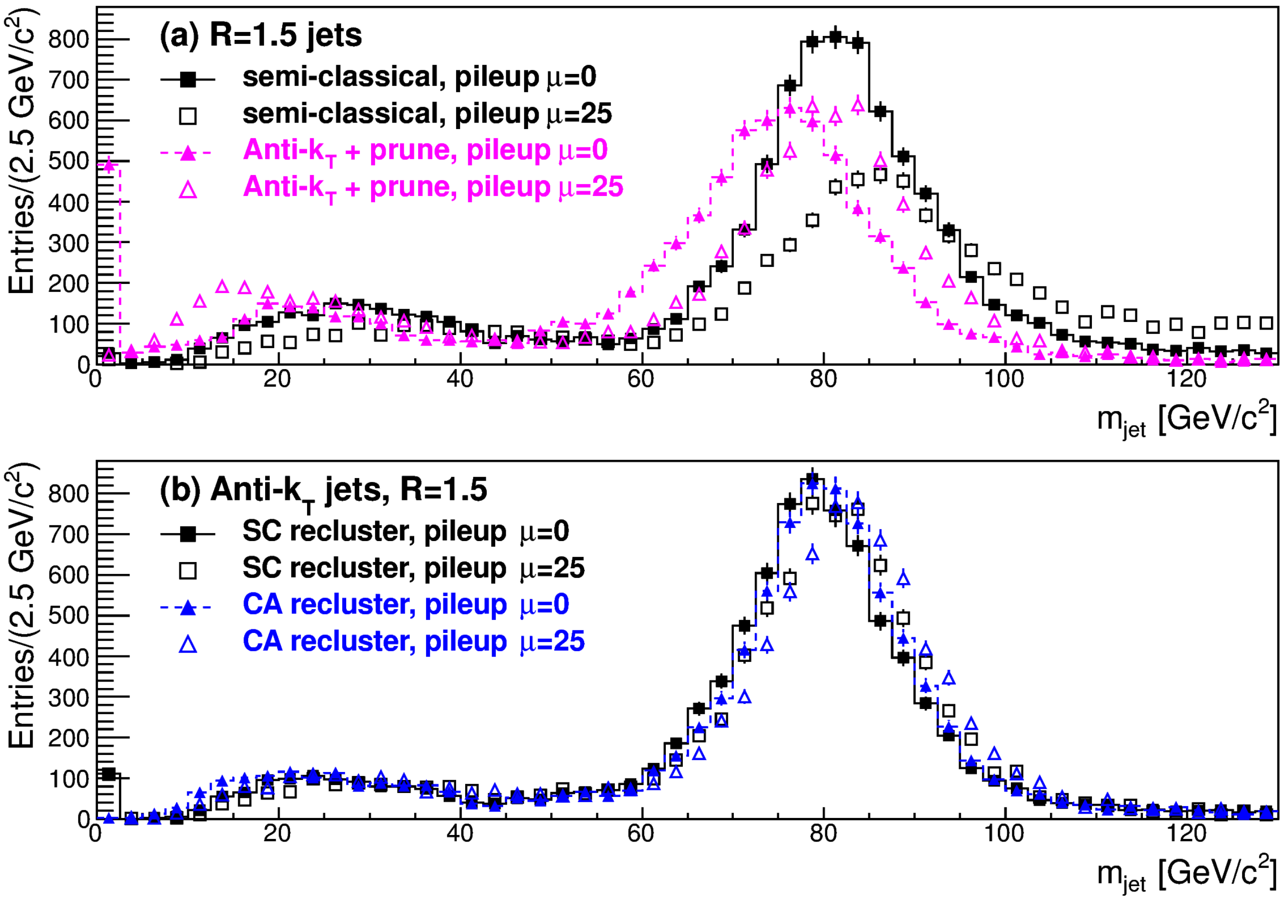
Stability vs jet size (2)

- More noticeable difference at higher pileup
  - Average 25 ~ current LHC

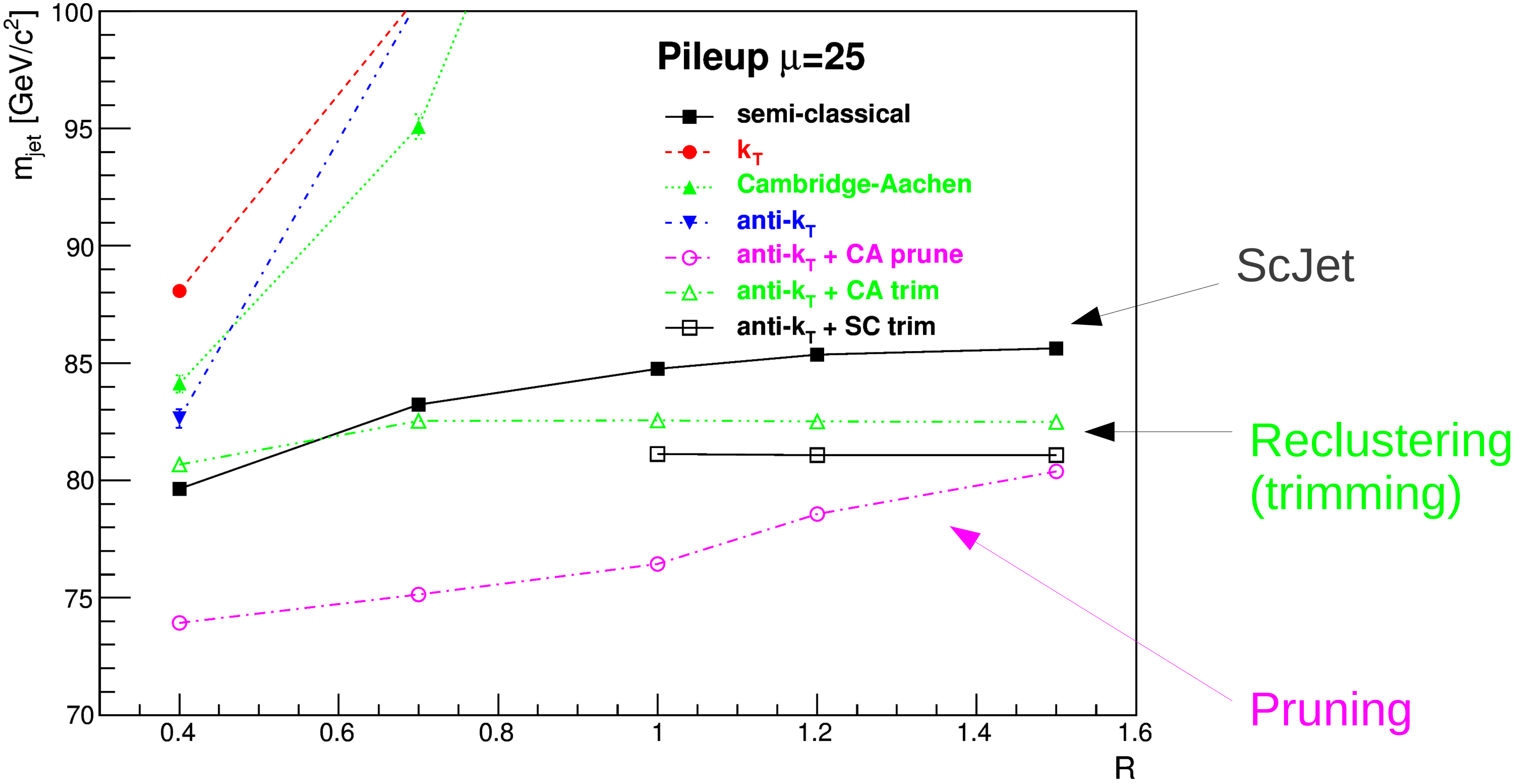


Comparison with grooming

- Pruning
  - $z_{cut} = 0.1$
  - $D_{cut} = 0.2$
- Trimming
  - $R_{filt} = 0.3$
  - $f_{sub} = 0.05$
  - Appears to be quite stable

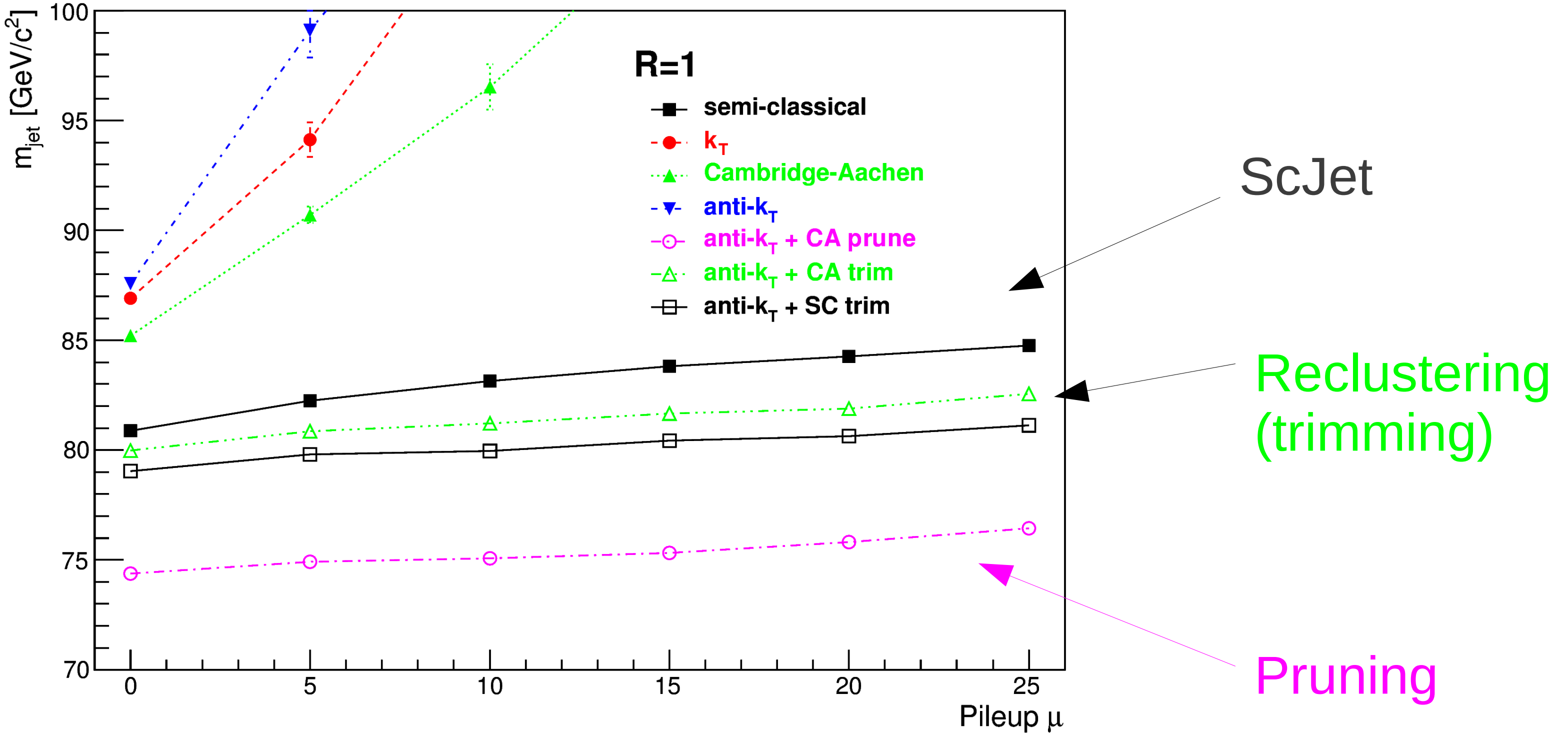


Grooming with pileup



Stability vs pileup

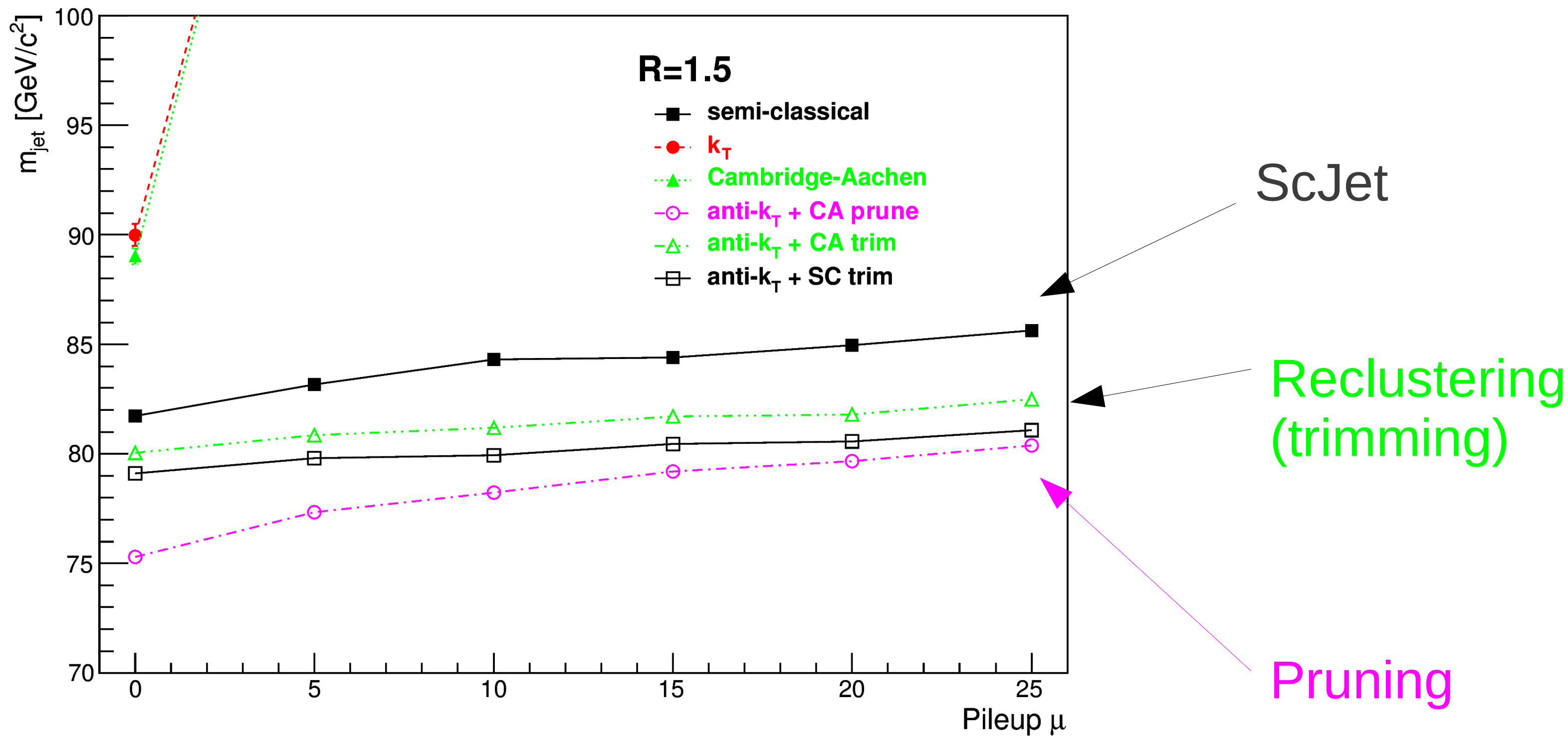
- W peak mass vs pileup level, *R*=1



- Again, ScJet stability ~ grooming

Stability vs pileup (2)

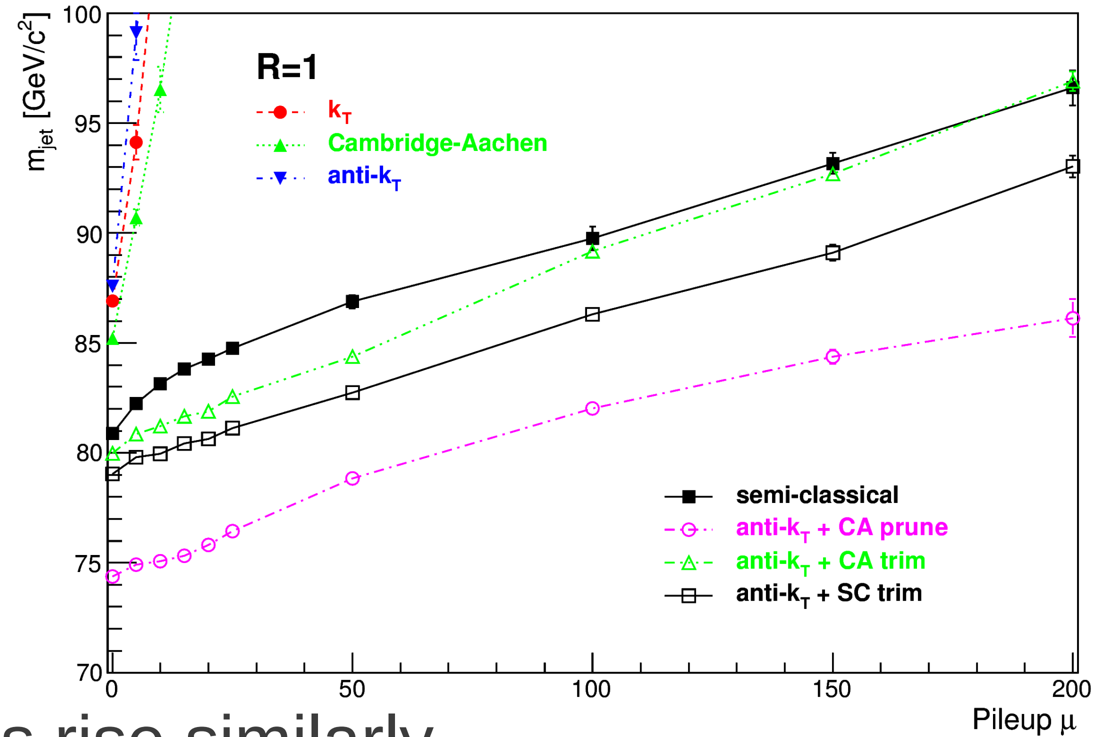
- With very fat jets, R=1.5



- ScJet and CA trimming appear most stable in this μ range
- At this point, it would be interesting to apply to LHC data
  - At the same time, look at some related issues

High luminosity

- Extend to HL-LHC pileup levels
  - Phase 1: ~25
  - Phase 1.5: 55-80
  - Phase 2: 140 (levelled)



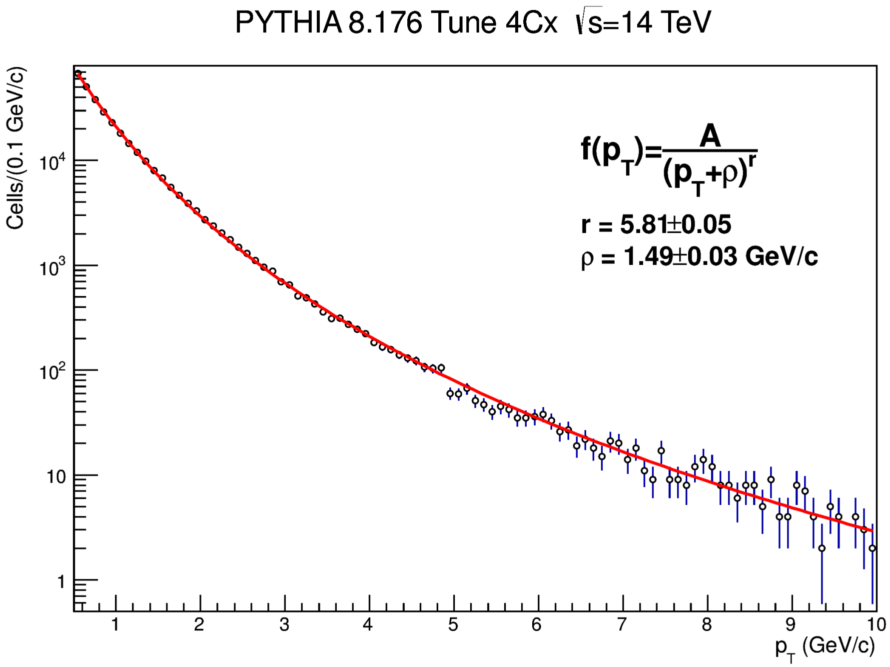
- All grooming techniques rise similarly
  - Need to tune for beam conditions – often subtract pileup
  - ScJet only has *R*, and zero jet area

Tuning ScJet

- Consider probabilities again: compare signal vs background
  - Signal: emission at some angle θ
  - Background: cell/cluster with some p_T
  - Larger μ → higher p_T
- d_{iB}: “beam-jet” distance, compare with inter-jet distance
  - Actually used in inclusive algorithms to limit ΔR_{ij} of mergings
  - Can it reflect an actual distance to compare with d_{ij}?
- ➢ Introduce a pileup scale to d_{iB}

Tuning ScJet (2)

Cell p_T's for 1 minbias



From cell p_T's with more pileup:

$$\rho(\mu)=(1.02+0.0117\mu)\,GeV$$

- Integrate with $F(p_T)=1$, flip

$$d_{iB}=\left(1+\frac{p_{Ti}}{k_{scale}\,\rho(\mu)}\right)^{1-r}$$

  - k_{scale} = scaling factor for ρ
  - Clustering reduces *r*
  - Take *r* = 5 and ρ(μ)
- With data, could use different d_{iB} for different background shape

“ScSubJet”

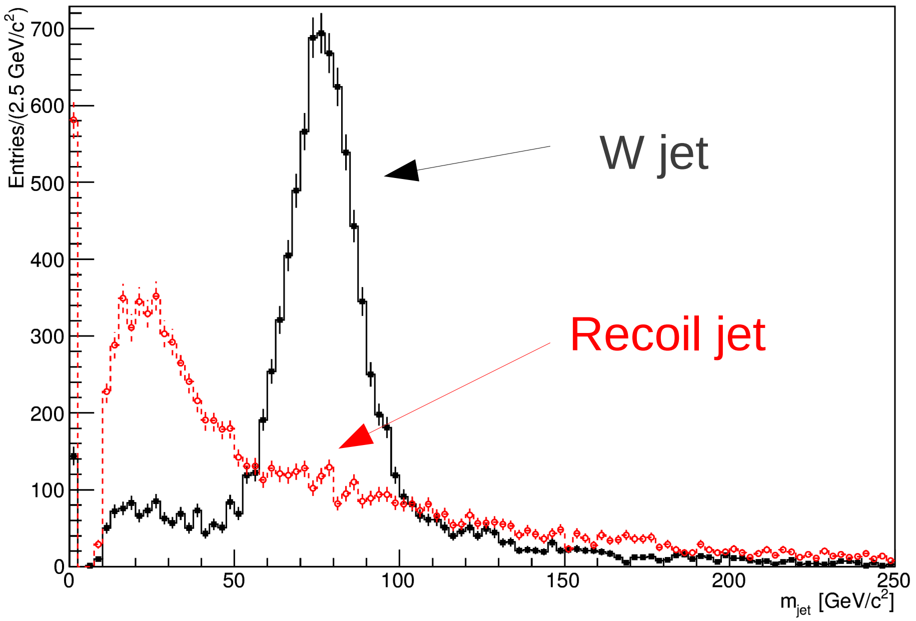
- Compare with probability of emission with angle > θ

$$d_{ij} = 1 + \gamma_{ij}^2 \left(\frac{\Delta R_{ij}}{R_{sc}} \right)^2 \qquad \gamma_{ij}^2 = 1 + \frac{|\vec{p}_{Ti} + \vec{p}_{Tj}|^2}{m_{ij}^2}$$

- Note that R_{sc} is no longer maximum ΔR_{ij}
  - Tends to cluster everything that isn't identified as background
  - Introduce a termination condition,
    or use to recluster (groom) an existing jet

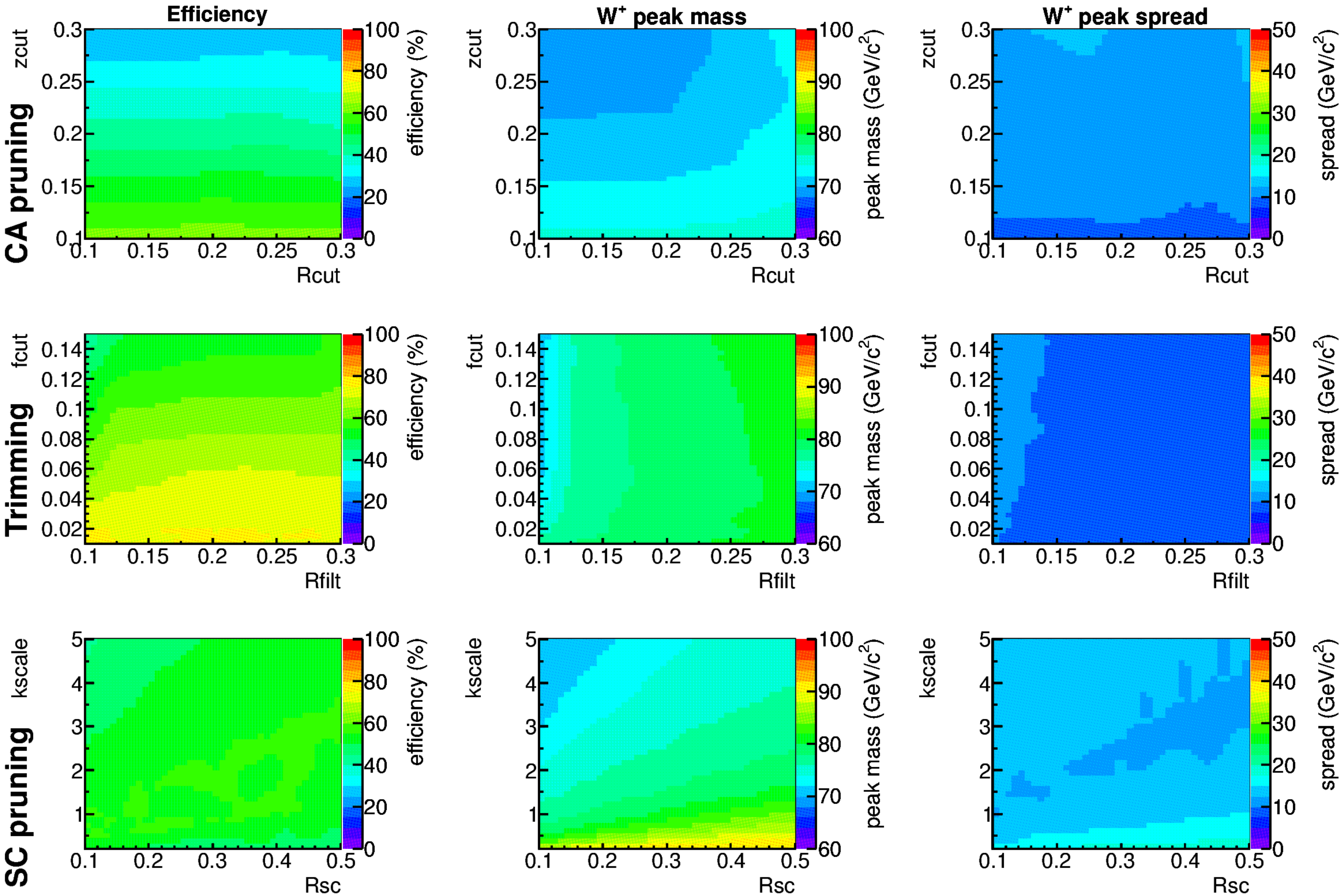
ScSubJet comparisons

- Start with anti-k_T jets with R=1
- Take R_{sc}=0.2, k_{scale}=1.5
- Look at W's with p_T > 500 GeV
- Also consider p_T > 160 GeV
  - 90% have daughter ΔR < 1
  - We don't know scale of new physics or signature



- Try to compare with good grooming parameters at different pileup and p_T ranges
  - Caveat: precision suitable for illustration only – still only simulation

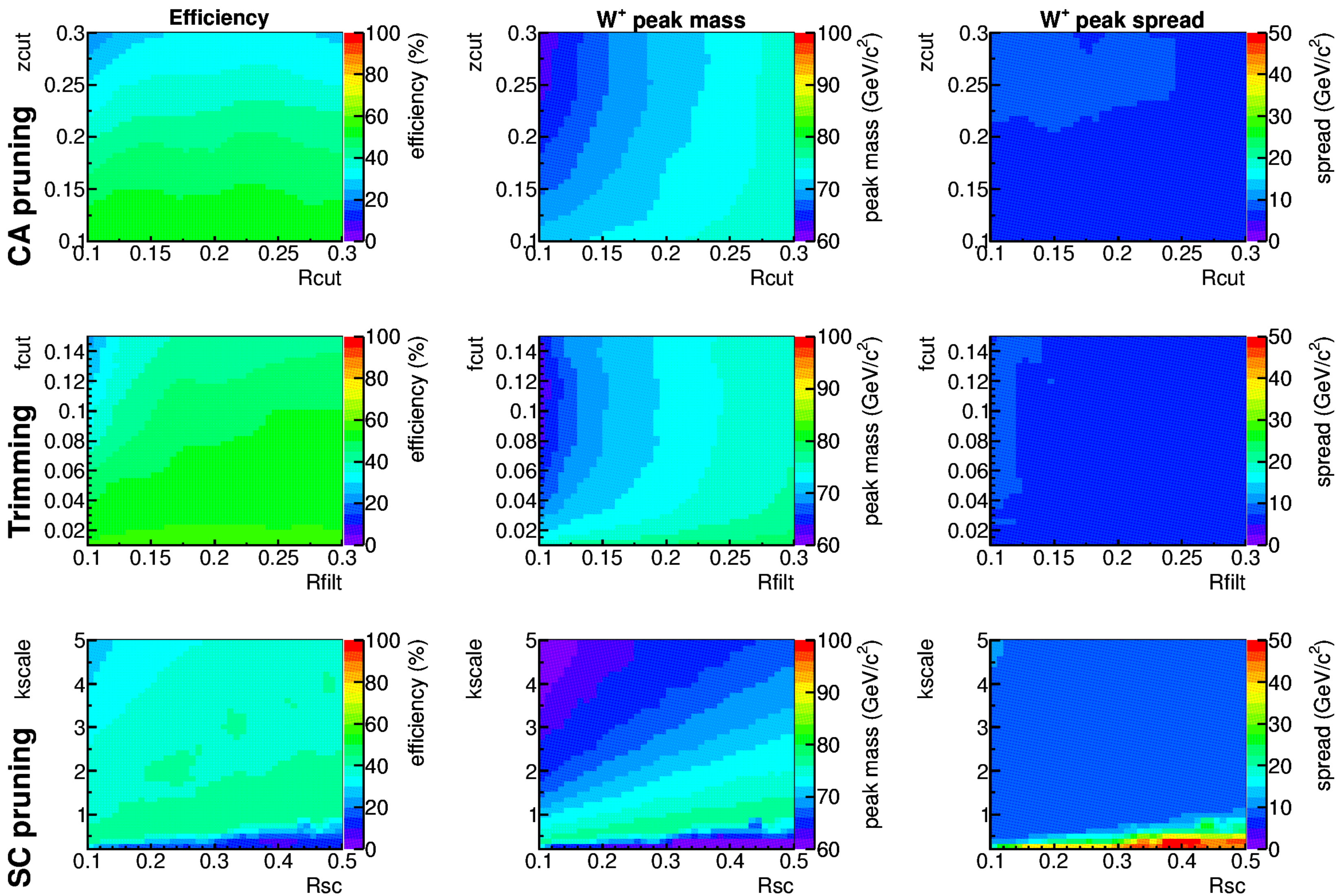
W p_T > 500 GeV, no pileup



Little bg → R does not make much difference

k_{scale} and R_{sc} complementary for most p_T's

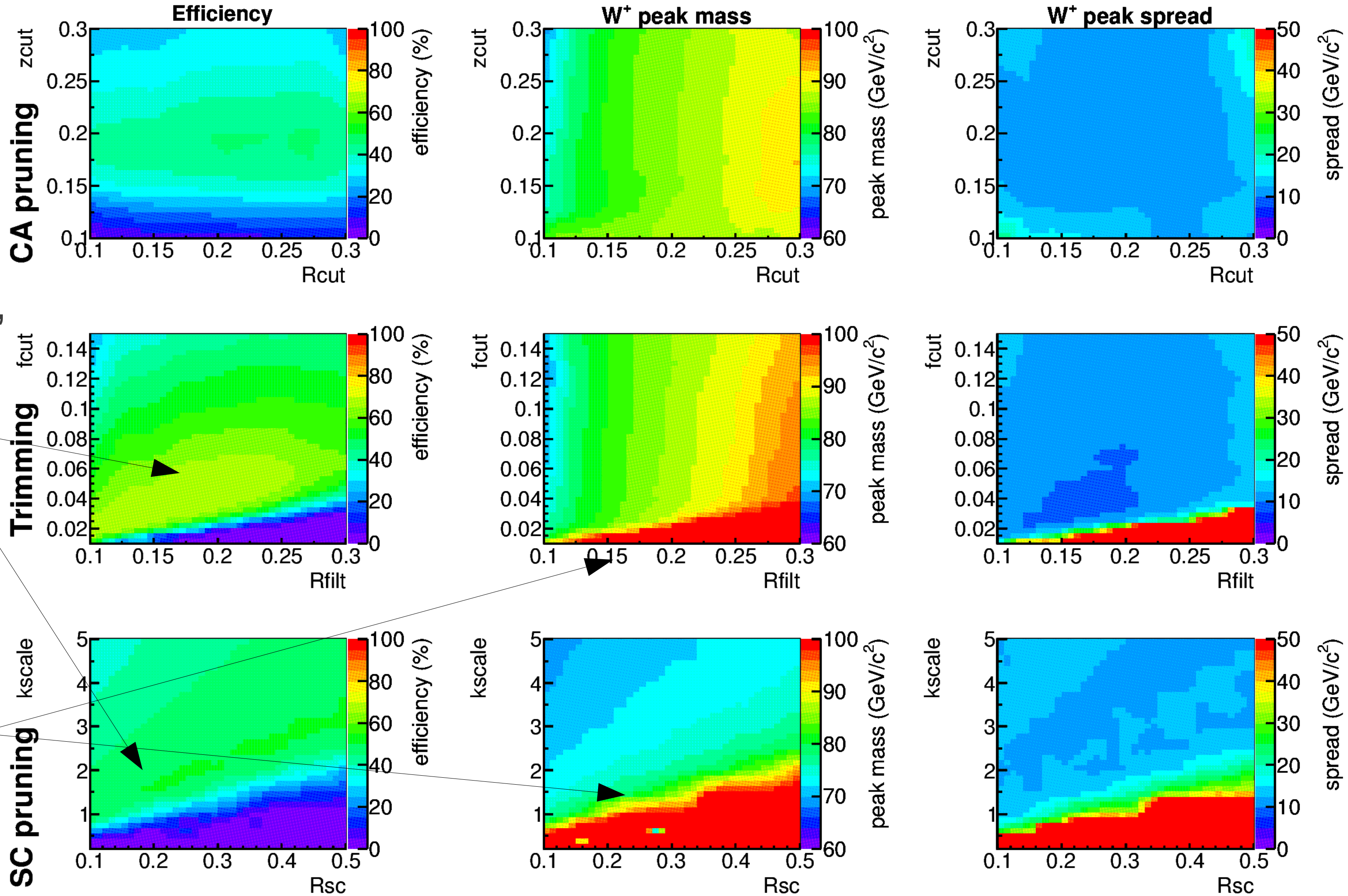
W p_T > 160 GeV, no pileup



More challenge: less boost

→ R cannot be too small

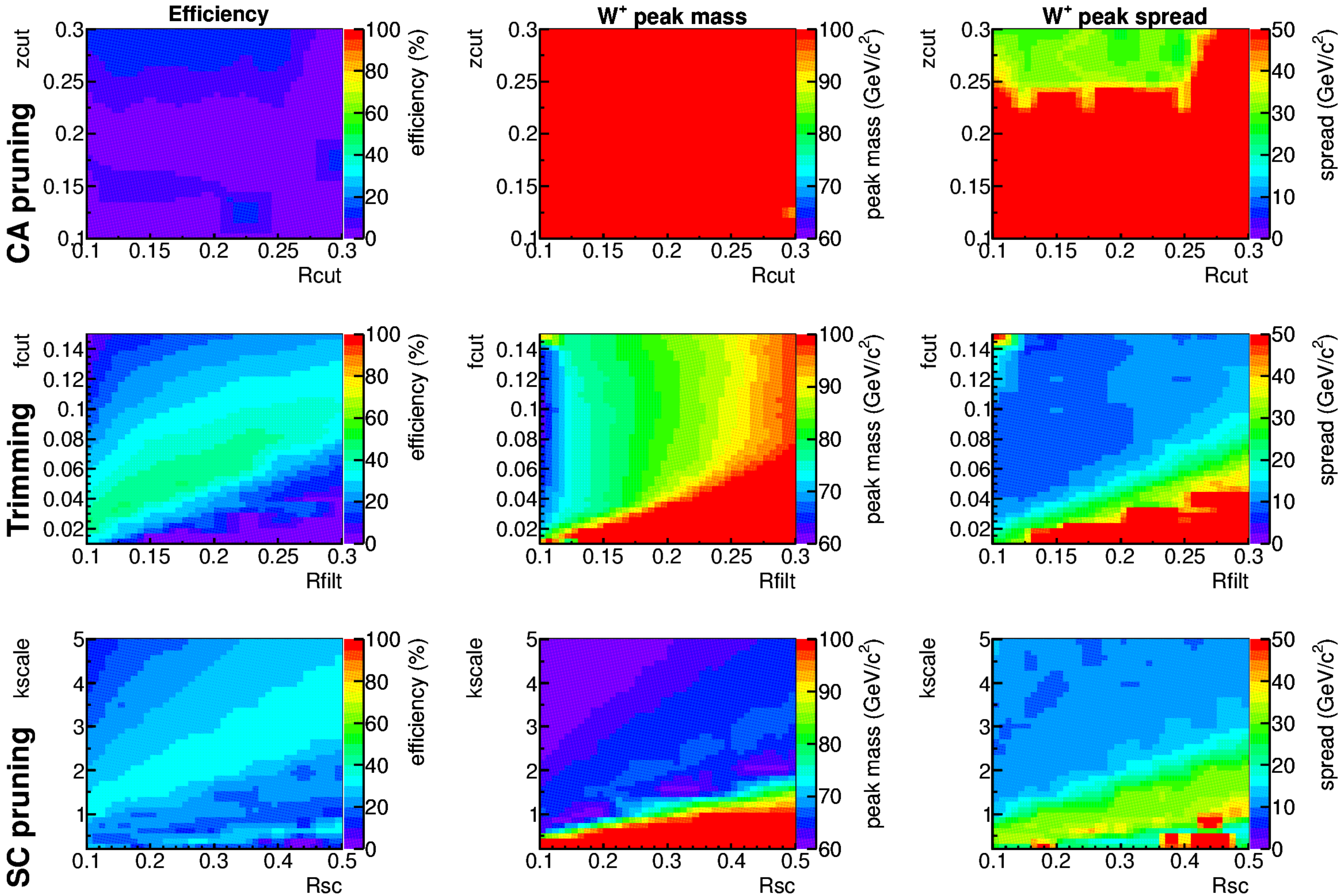
W p_T > 500 GeV, 150 pileup



Prefer smallish R, lower grooming power

But not so little grooming power that grooming stops working

W p_T > 160 GeV, 150 pileup

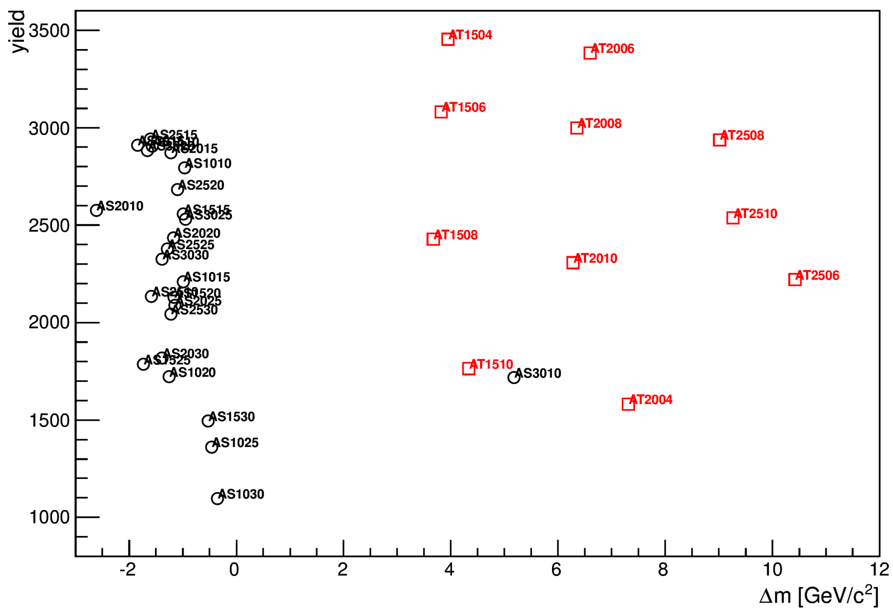


Most challenging: grooming stops working for many parameters

Still prefer smallish R and low grooming power

Parameter selection

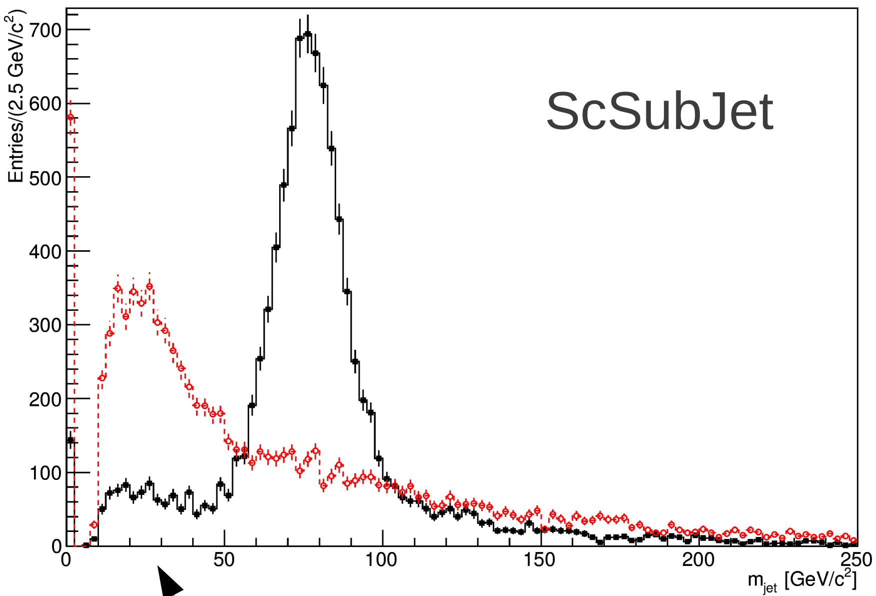
- Choose parameters with high yield, low Δm
  - Δm: peak mass drift from pileup μ → μ/2
- Good parameters for both μ=50 and μ=200
- Trimming:
  R_{filt} = 0.15
  f_{sub} = 0.04
- ScSubJet:
  R_{sc} = 0.2
  k_{scale} = 1.5

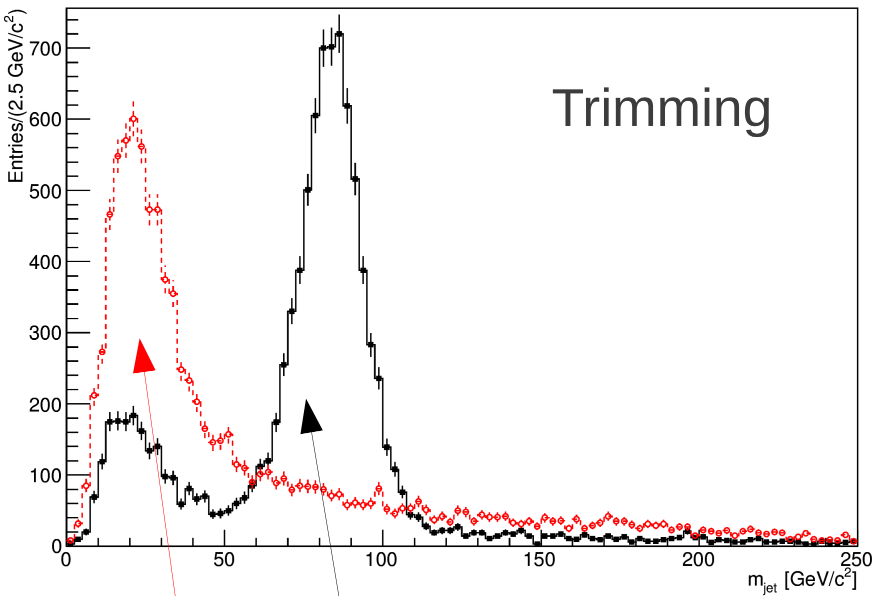


W p_T > 500 GeV
200 pileup

Comparison with parameters

W p_T > 500 GeV, 200 pileup



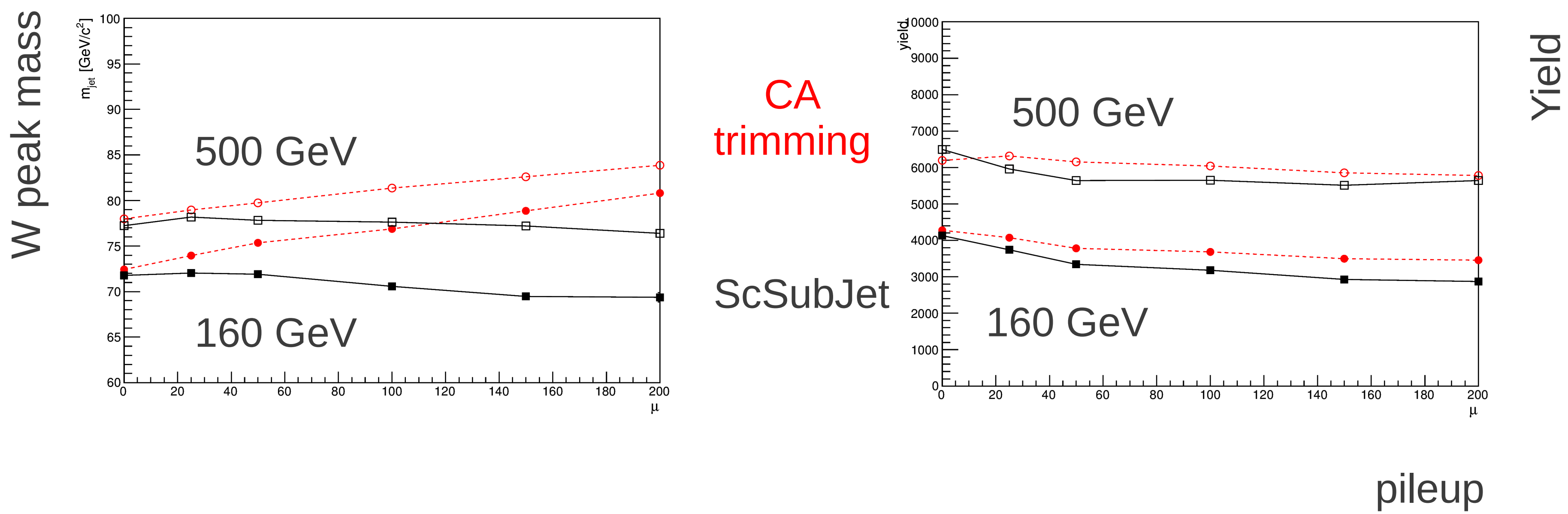


Reduced “overgrooming”

Slightly narrower peak

Narrower recoil jet mass

Comparison with parameters (2)

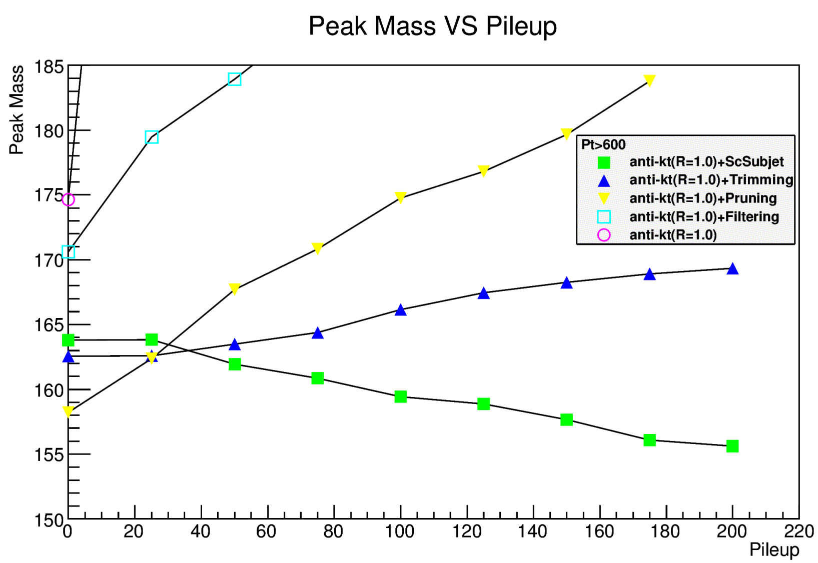


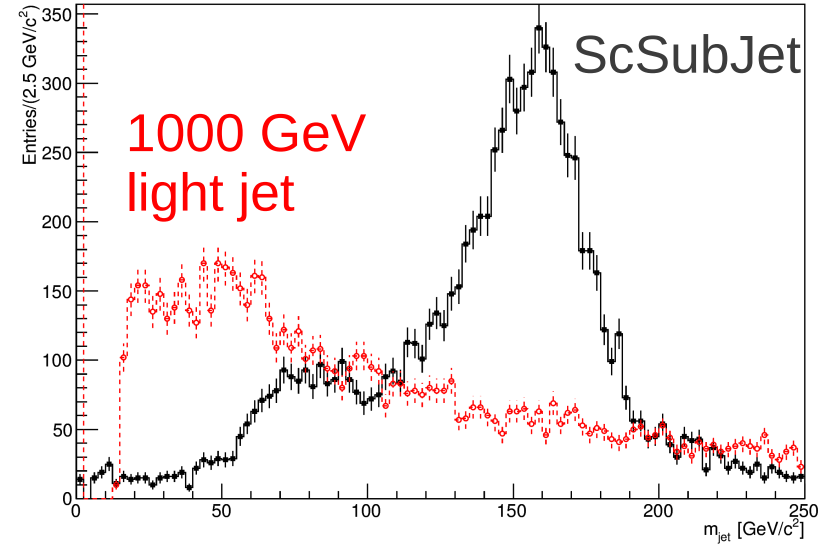
- Trimming W peak still drifts upwards, but overall slightly more efficient
- ScSubJet peak descends slightly
  - Possibly overshot restoring stability via d_{iB}: still room for tuning ρ(μ) and d_{iB}
- Concern: p_T-dependent W mass
  - Smears peak → more background for subsequent tagging
  - Not like pileup level, over which physicists have some control

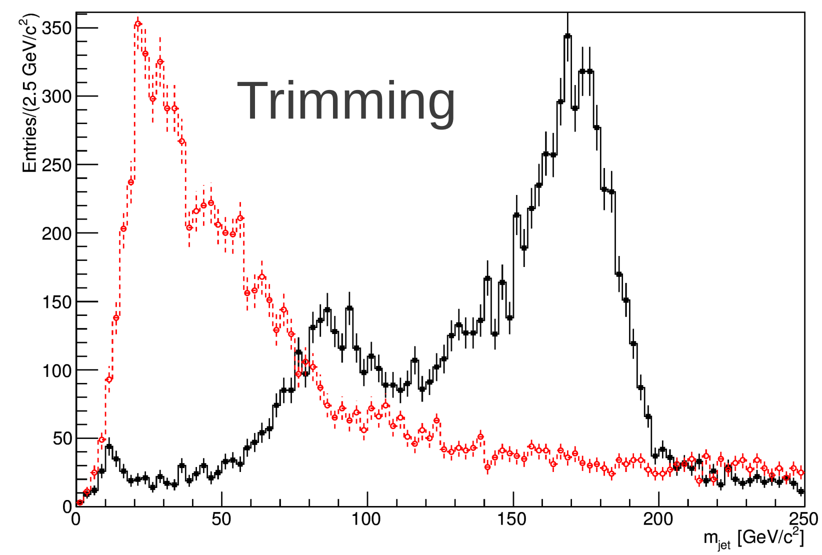
Top quarks

$Z' \rightarrow t\bar{t},\ m = 2\,TeV$

- Initial tests using selected parameters from W:
  - Similar shape to other groomed distributions
  - Similar patterns in peak mass, width





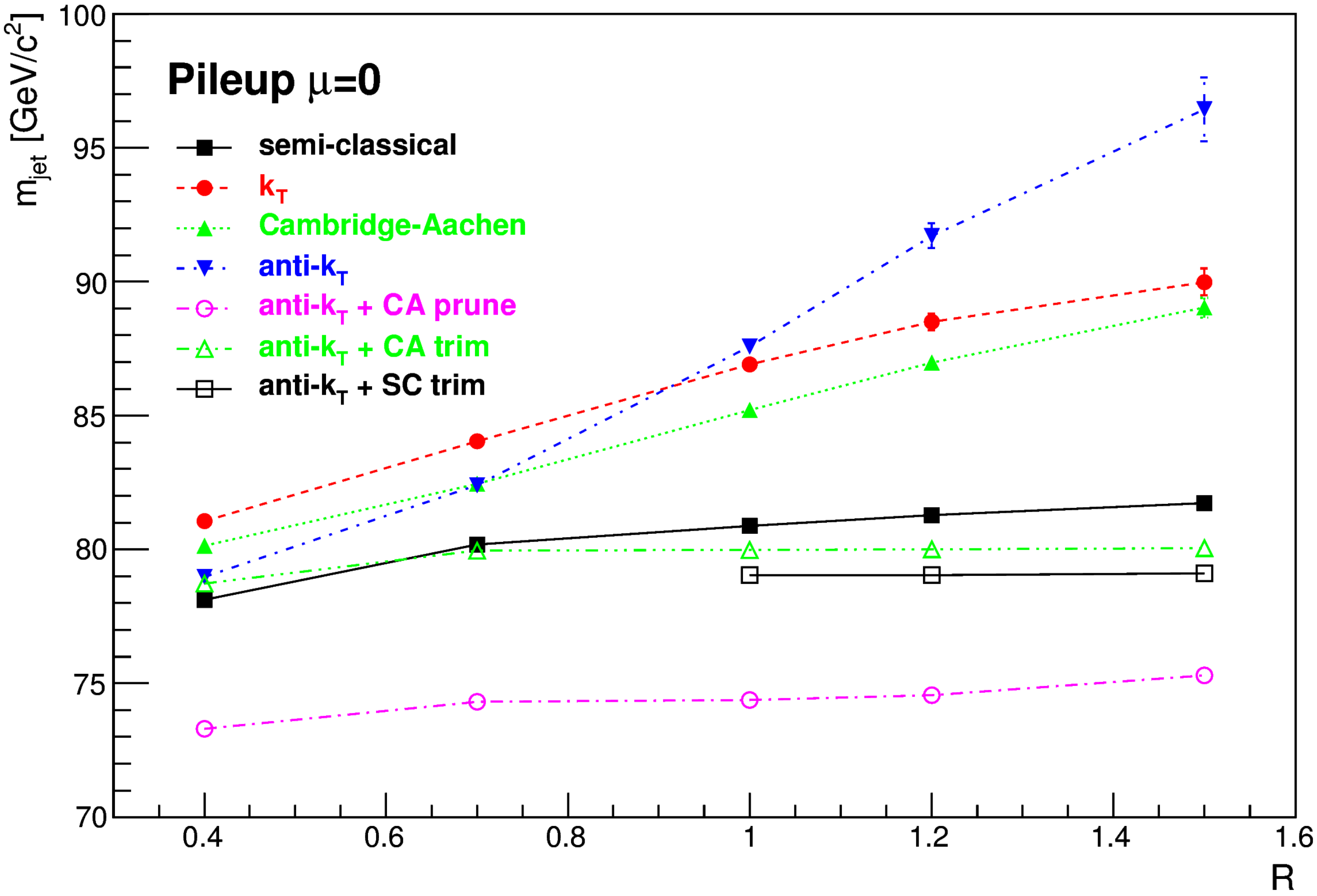


Conclusion

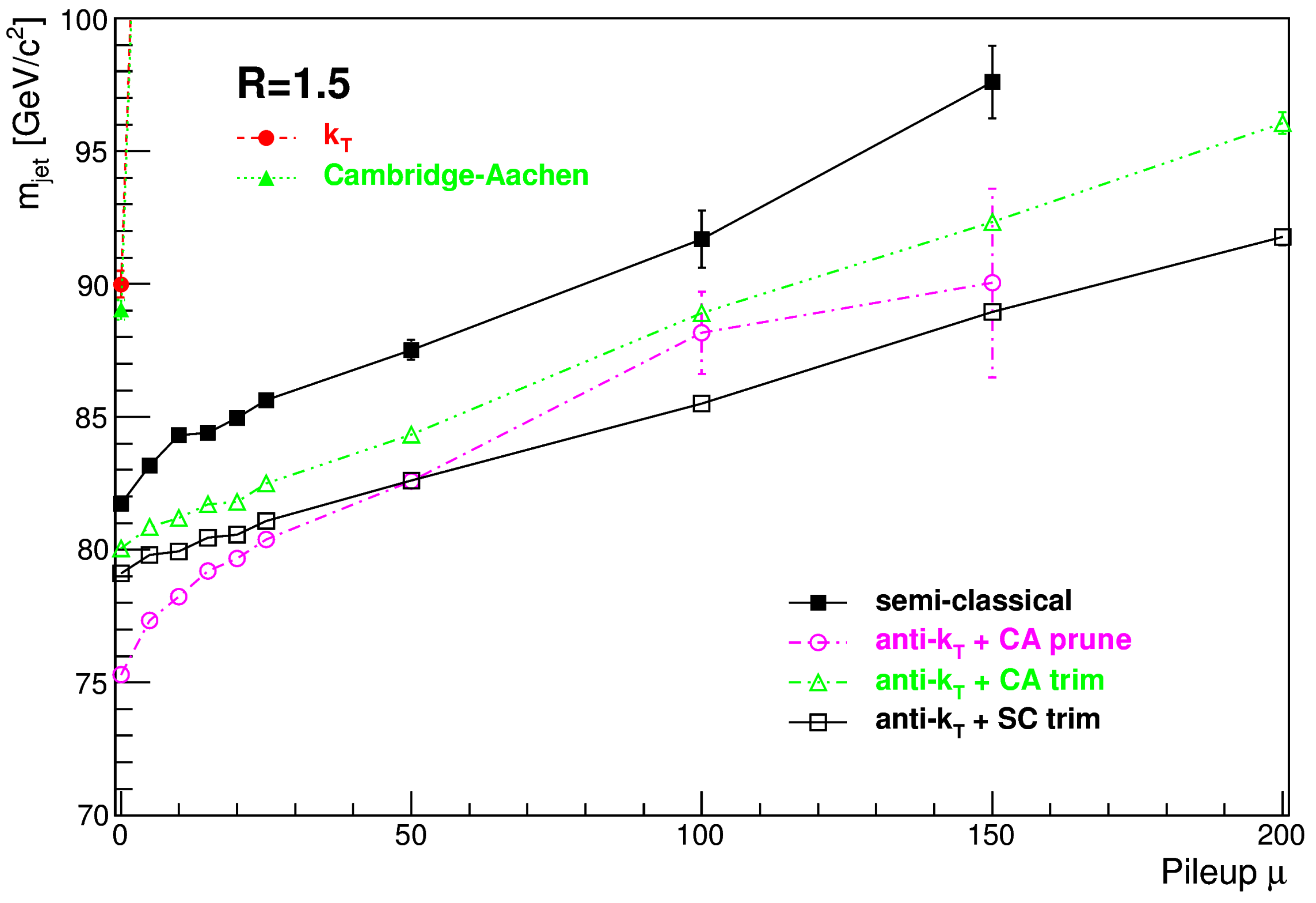
- ScJet clusters and grooms simultaneously
  - Even for reclustering, peak mass drifts, but fewer handles to turn
- ScSubJet incorporates pileup-dependent p_T scale into d_{iB}
  - Rough “tuning” is more stable in W peak mass than best trimming
  - Slightly lower yield and wider peak in Pythia8 MC
  - Current background model has room for improvement
  - Starting to look at whether ScSubJet approach can help improve top tagging and other jet observables
- There may be some mileage in reexamining distance measures in jet clustering algorithms
  - Considering background model appears to have some benefit
  - ScSubJet for reclustering can be seen as another form of pruning, with performance which can be tuned to be comparable to that of trimming

Backup

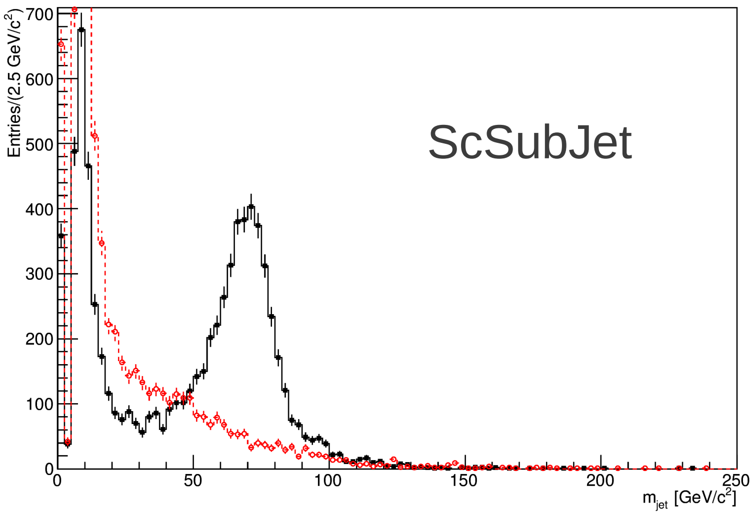
Grooming without pileup

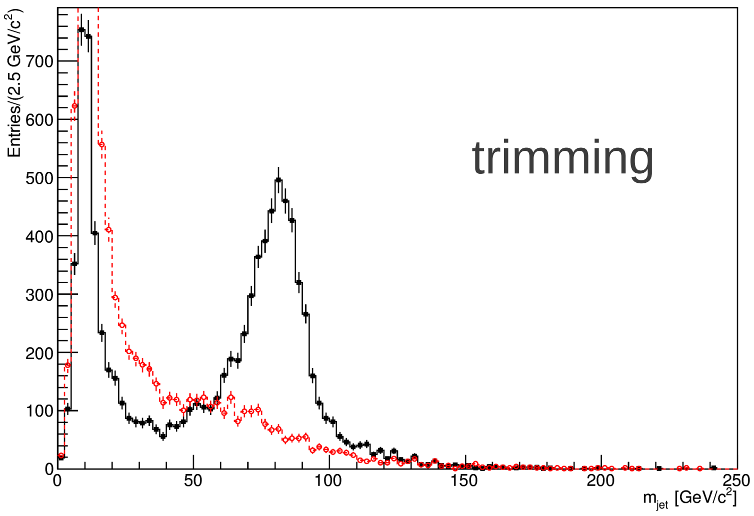


High luminosity, fat jets



Comparison with parameters (3)





- W p_T > 160 GeV, pileup 200